

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
9 August 2001 (09.08.2001)

PCT

(10) International Publication Number
WO 01/57276 A2

(51) International Patent Classification⁷: **C12Q 1/68**

94043 (US). RANK, David, R. [US/US]; 117 El Dorado Commons, Fremont, CA 94539 (US).

(21) International Application Number: PCT/US01/00668

(74) Agent: RONNING, Royal, N., Jr.; Amersham Pharmacia Biotech, Inc., 800 Centennial Avenue, Piscataway, NJ 08855 (US).

(22) International Filing Date: 30 January 2001 (30.01.2001)

(25) Filing Language: English

(81) Designated States (*national*): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(26) Publication Language: English

(30) Priority Data:
60/180,312 4 February 2000 (04.02.2000) US
60/207,456 26 May 2000 (26.05.2000) US
09/608,408 30 June 2000 (30.06.2000) US
09/632,366 3 August 2000 (03.08.2000) US
60/234,687 21 September 2000 (21.09.2000) US
60/236,359 27 September 2000 (27.09.2000) US
0024263.6 4 October 2000 (04.10.2000) GB

(84) Designated States (*regional*): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

(71) Applicant (*for all designated States except US*): MOLECULAR DYNAMICS, INC. [—/US]; 928 East Arques Avenue, Sunnyvale, CA 94086 (US).

Published:

- without international search report and to be republished upon receipt of that report
- entirely in electronic form (except for this front page) and available upon request from the International Bureau

(72) Inventors; and

(75) Inventors/Applicants (*for US only*): PENN, Sharron, G. [GB/US]; 617 South Delaware Street, San Mateo, CA 94402 (US). HANZEL, David, K. [US/US]; 988 Loma Verde Avenue, Palo Alto, CA 94303 (US). CHEN, Wensheng [CN/US]; 210 Easy Street #25, Mountain View, CA

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR ANALYSIS OF GENE EXPRESSION IN HUMAN BONE MARROW

(57) Abstract: A single exon nucleic acid microarray comprising a plurality of single exon nucleic acid probes for measuring gene expression in a sample derived from human bone marrow is described. Also described are single exon nucleic acid probes expressed in the bone marrow and their use in methods for detecting gene expression.

WO 01/57276 A2

HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL
FOR ANALYSIS OF GENE EXPRESSION IN HUMAN BONE MARROW

CROSS REFERENCE TO RELATED APPLICATIONS

5

The present application is a continuation-in-part of U.S.
patent application serial nos. 09/632,366, filed August 3,
2000 and 09/608,408, filed June 30, 2000; claims the
benefit under 35 U.S.C. s 119(e) of U.S. provisional patent
10 application serial nos. 60/236,359, filed September 27,
2000, 60/234,687, filed September 21, 2000, 60/207,456,
filed May 26, 2000, and 60/180,312, filed February 4, 2000;
and further claims the benefit under 35 U.S.C. s 119(a) of
UK patent application no. 0024263.6, filed October 4, 2000,
15 the disclosures of which are incorporated herein by
reference in their entireties.

REFERENCE TO SEQUENCE LISTING AND INCORPORATION BY
REFERENCE THEREOF

20

The present application includes a Sequence Listing in
electronic format, filed pursuant to PCT Administrative
Instructions 801 - 806 on a single CD-R disc, in
triplicate, containing a file named pto_BONE_MARROW.txt,
25 created 24 January 2001, having 26,421,347 bytes. The
Sequence Listing contained in said file on said disc is
incorporated herein by reference in its entirety.

Field of the Invention

30

The present invention relates to genome-derived
single exon microarrays useful for verifying the expression
of regions of genomic DNA predicted to encode protein. In
particular, the present invention relates to unique genome-
35 derived single exon nucleic acid probes expressed in human

bone marrow and single exon nucleic acid microarrays that include such probes.

Background of the Invention

5 For almost two decades following the invention of general techniques for nucleic acid sequencing, Sanger et al., *Proc. Natl. Acad. Sci. USA* 70(4):1209-13 (1973); Gilbert et al., *Proc. Natl. Acad. Sci. USA* 70(12):3581-4 (1973), these techniques were used principally as tools to
10 further the understanding of proteins — known or suspected — about which a basic foundation of biological knowledge had already been built. In many cases, the cloning effort that preceded sequence identification had been both informed and directed by that antecedent
15 biological understanding.

 For example, the cloning of the T cell receptor for antigen was predicated upon its known or suspected cell type-specific expression, by its suspected membrane association, and by the predicted assembly of its gene via
20 T cell-specific somatic recombination. Subsequent sequencing efforts at once confirmed and extended understanding of this family of proteins. Hedrick et al., *Nature* 308(5955):153-8 (1984).

 More recently, however, the development of high
25 throughput sequencing methods and devices, in concert with large public and private undertakings to sequence the human and other genomes, has altered this investigational paradigm: today, sequence information often precedes understanding of the basic biology of the encoded protein
30 product.

 One of the approaches to large-scale sequencing is predicated upon the proposition that expressed sequences — that is, those accessible through isolation of mRNA — are of greatest initial interest. This "expressed
35 sequence tag" ("EST") approach has already yielded vast

amounts of sequence data (see for example Adams et al.,
Science 252:1651 (1991); Williamson, *Drug Discov. Today*
4:115 (1999)). For nucleic acids sequenced by this
approach, often the only biological information that is
5 known *a priori* with any certainty is the likelihood of
biologic expression itself. By virtue of the species and
tissue from which the mRNA had originally been obtained,
most such sequences are also annotated with the identity of
the species and at least one tissue in which expression
10 appears likely.

More recently, the pace of genomic sequencing has
accelerated dramatically. When genomic DNA serves as the
initial substrate for sequencing efforts, expression cannot
be presumed; often the only *a priori* biological information
15 about the sequence includes the species and chromosome (and
perhaps chromosomal map location) of origin.

With the ever-accelerating pace of sequence
accumulation by directed, EST, and genomic sequencing
approaches – and in particular, with the accumulation of
20 sequence information from multiple genera, from multiple
species within genera, and from multiple individuals within
a species – there is an increasing need for methods that
rapidly and effectively permit the functions of nucleic
sequences to be elucidated. And as such functional
25 information accumulates, there is a further need for
methods of storing such functional information in
meaningful and useful relationship to the sequence itself;
that is, there is an increasing need for means and
apparatus for annotating raw sequence data with known or
30 predicted functional information.

Although the increase in the pace of genomic
sequencing is due in large part to technological changes in
sequencing strategies and instrumentation, Service, *Science*
280:995 (1998); Pennisi, *Science* 283: 1822-1823 (1999),
35 there is an important functional motivation as well.

While it was understood that the EST approach would rarely be able to yield sequence information about the noncoding portions of the genome, it now also appears the EST approach is capable of capturing only a fraction of
5 a genome's actual expression complexity.

For example, when the *C. elegans* genome was fully sequenced, gene prediction algorithms identified over 19,000 potential genes, of which only 7,000 had been found by EST sequencing. *C. elegans* Sequencing Consortium,
10 *Science* 282:2012 (1998). Analogously, the recently completed sequence of chromosome 2 of *Arabidopsis* predicts over 4000 genes, Lin et al., *Nature*, 402:761 (1999), of which only about 6% had previously been identified via EST sequencing efforts. Although the human genome has the
15 greatest depth of EST coverage, it is still woefully short of surrendering all of its genes. One recent estimate suggests that the human genome contains more than 146,000 genes, which would at this point leave greater than half of the genes undiscovered. It is now predicted that many
20 genes, perhaps 20 to 50%, will only be found by genomic sequencing.

There is, therefore, a need for methods that permit the functional regions of genomic sequence – and most importantly, but not exclusively, regions that
25 function to encode genes – to be identified.

Much of the coding sequence of the human genome is not homologous to known genes, making detection of open reading frames ("ORFs") and predictions of gene function difficult. Computational methods exist for predicting
30 coding regions in eukaryotic genomes. Gene prediction programs such as GRAIL and GRAIL II, Uberbacher et al., *Proc. Natl. Acad. Sci. USA* 88(24):11261-5 (1991); Xu et al., *Genet. Eng.* 16:241-53 (1994); Uberbacher et al., *Methods Enzymol.* 266:259-81 (1996); GENEFINDER, Solovyev et
35 al., *Nucl. Acids. Res.* 22:5156-63 (1994); Solovyev et al.,

Ismb 5:294-302 (1997); and GENESCAN, Burge et al., *J. Mol. Biol.* 268:78-94 (1997), predict many putative genes without known homology or function. Such programs are known, however, to give high false positive rates. Burset et al.,
5 *Genomics* 34:353-367 (1996). Using a consensus obtained by a plurality of such programs is known to increase the reliability of calling exons from genomic sequence. Ansari-Lari et al., *Genome Res.* 8(1):29-40 (1998)

Identification of functional genes from genomic
10 data remains, however, an imperfect art. For example, in reporting the full sequence of human chromosome 21, the Chromosome 21 Mapping and Sequencing Consortium reports that prior bioinformatic estimates of human gene number may need to be revised substantially downwards. *Nature*
15 405:311-199 (2000); Reeves, *Nature* 405:283-284 (2000).

Thus, there is a need for methods and apparatus that permit the functions of the regions identified bioinformatically - and specifically, that permit the expression of regions predicted to encode protein - readily
20 to be confirmed experimentally.

Recently, the development of nucleic acid microarrays has made possible the automated and highly parallel measurement of gene expression. Reviewed in Schena (ed.), DNA Microarrays : A Practical Approach
25 (Practical Approach Series), Oxford University Press (1999) (ISBN: 0199637768); *Nature Genet.* 21(1)(suppl):1 - 60 (1999); Schena (ed.), Microarray Biochip: Tools and Technology, Eaton Publishing Company/BioTechniques Books Division (2000) (ISBN: 1881299376).

30 It is common for microarrays to be derived from cDNA/EST libraries, either from those previously described in the literature, such as those from the I.M.A.G.E. consortium, Lennon et al., *Genomics* 33(1):151-2 (1996), or from the construction of "problem specific" libraries
35 targeted at a particular biological question, R.S. Thomas

et al., *Cancer Res.* (in press). Such microarrays by definition can measure expression only of those genes found in EST libraries, and thus have not been useful as probes for genes discovered solely by genomic sequencing.

5 The utility of using whole genome nucleic acid microarrays to answer certain biological questions has been demonstrated for the yeast *Saccharomyces cerevisiae*. De Risi et al., *Science* 278:680 (1997). The vast majority of yeast nuclear genes, approximately 95% however, are single
10 exon genes, i.e., lack introns, Lopez et al., *RNA* 5:1135-1137 (1999); Goffeau et al., *Science* 274:563-67 (1996), permitting coding regions more readily to be identified. Whole genome nucleic acid microarrays have not generally been used to probe gene expression from more complex
15 eukaryotic genomes, and in particular from those averaging more than one intron per gene.

 Because bone marrow is the tissue in which blood cells originate, diseases of the bone marrow are a significant cause of human morbidity and mortality.
20 Increasingly, genetic factors are being found that contribute to predisposition, onset, and/or aggressiveness of most, if not all, of these diseases. Although mutations in single genes have in some cases been identified as causal - notably in the thalassemias and sickle cell anemia
25 - disorders of the bone marrow are, for the most part, believed to have polygenic etiologies. There is a need for methods and apparatus that permit prediction, diagnosis and prognosis of diseases of the bone marrow, particularly those diseases with polygenic etiology.

30

Summary of the Invention

 The present invention solves these and other problems in the art by providing methods and apparatus for
35 predicting, confirming, and displaying functional

information derived from genomic sequence. The present invention also provides apparatus for verifying the expression of putative genes identified within genomic sequence.

5 In particular, the invention provides novel genome-derived single exon nucleic acid microarrays useful for verifying the expression of putative genes identified within genomic sequence.

 The present invention also provides compositions
10 and kits for the ready production of nucleic acids identical in sequence to, or substantially identical in sequence to, probes on the genome-derived single exon microarrays of the present invention.

 Accordingly, in a first aspect of the invention,
15 there is provided a spatially-addressable set of single exon nucleic acid probes for measuring gene expression in a sample derived from human bone marrow, comprising a plurality of single exon nucleic acid probes according to any one of the nucleotide sequences set out in SEQ ID NOs:
20 1 - 13,114 or a complementary sequence, or a portion of such a sequence.

 By plurality is meant at least two, suitably at least 20, most suitably at least 100, preferably at least 1000 and, most preferably, upto 5000.

25 In one embodiment of the first aspect, each of said plurality of probes is separately and addressably amplifiable.

 In an alternative embodiment, each of said plurality of probes is separately and addressably
30 isolatable from said plurality.

 In a preferred embodiment, each of said plurality of probes is amplifiable using at least one common primer. Preferably, each of said plurality of probes is amplifiable using a first and a second common primer.

35 In yet another embodiment, said set of single

exon nucleic acid probes comprises between 50 - 20,000 probes, for example, 50 - 5000.

Suitably, said set of single exon nucleic acid probes comprises at least 50 - 1000 discrete single exon nucleic acid probes having a sequence as set out in any of SEQ ID NOS.: 1 - 26,012 or a complimentary sequence, or a portion of such a sequence.

Preferably, the average length of the single exon nucleic acid probes is between 200 and 500 bp. It is preferred that the average length should be at least 200bp, suitably at least 250bp, most suitably at least 300bp, preferably at least 400bp and, most preferably, 500 bp.

In another embodiment, the single exon nucleic acid probes lack prokaryotic and bacteriophage vector sequence. It is preferred that at least 50%, suitably at least 60%, most suitably at least 70%, preferably at least 75%, more preferably at least 80, 85, 90, 95 or 99% of said single exon nucleic acid probes lack prokaryotic and bacteriophage vector sequence.

In another preferred embodiment, said single exon nucleic acid lack homopolymeric stretches of A or T. It is preferred that at least 50%, suitably at least 60%, most suitably at least 70%, preferably at least 75%, more preferably at least 80, 85, 90, 95 or 99% of said single exon nucleic acid probes lack homopolymeric stretches of A or T.

Preferably, a spatially-addressable set of single exon nucleic acid probes in accordance with the first aspect of the invention is addressably disposed upon a substrate.

Suitable substrates include a filter membrane which may, preferably, be nitrocellulose or nylon. The nylon may preferably, be positively-charged. Other suitable substrates include glass, amorphous silicon, crystalline silicon, and plastic. Further suitable materials include

polymethylacrylic, polyethylene, polypropylene,
polyacrylate, polymethylmethacrylate, polyvinylchloride,
polytetrafluoroethylene, polystyrene, polycarbonate,
polyacetal, polysulfone, celluloseacetate,
5 cellulosenitrate, nitrocellulose, and mixtures thereof.

In a second aspect of the invention, there is
provided a microarray comprising a spatially addressable
set of single exon nucleic acid probes in accordance with
the first aspect of the invention.

10 In one embodiment, a genome-derived single-exon
microarray is packaged together with such an ordered set of
amplifiable probes corresponding to the probes, or one or
more subsets of probes, thereon. In alternative
embodiments, the ordered set of amplifiable probes is
15 packaged separately from the genome-derived single exon
microarray.

In another aspect, the invention provides genome-
derived single exon nucleic acid probes useful for gene
expression analysis, and particularly for gene expression
20 analysis by microarray. In particular embodiments of this
aspect, the present invention provides human single-exon
probes that include specifically-hybridizable fragments of
SEQ ID Nos. 13,115 - 26,012, wherein the fragment
hybridizes at high stringency to an expressed human gene.
25 In particular embodiments, the invention provides single
exon probes comprising SEQ ID Nos. 1 - 13,114.

Accordingly, in a third aspect of the invention,
there is provided a single exon nucleic acid probe for
measuring human gene expression in a sample derived from
30 human bone marrow which is a nucleic acid molecule
comprising a nucleotide sequence as set out in any of SEQ
ID NOS.: 1 - 13,114 or a complementary sequence or a
fragment thereof wherein said probe hybridizes at high
stringency to a nucleic acid expressed in the human bone
35 marrow.

In one embodiment, a single exon nucleic acid probe in accordance with the third aspect comprises a nucleotide sequence as set out in any of SEQ ID NOs.: 13,115 - 26,012 or a complementary sequence or a fragment thereof.

In a fourth aspect of the invention, there is provided a single exon nucleic acid probe for measuring human gene expression in a sample derived from human bone marrow which is a nucleic acid molecule having a sequence encoding a peptide comprising a peptide sequence as set out in any of SEQ ID NOs.: 26,013 - 38,628 or a complementary sequence or a fragment thereof wherein said probe hybridizes at high stringency to a nucleic acid expressed in the human bone marrow.

Preferably, a single exon nucleic acid probe in accordance with the third or fourth aspects of the invention comprises between at least 15 and 50 contiguous nucleotides of said SEQ ID NO:. It is preferred that the single exon nucleic acid probe comprises at least 15, suitably at least 20, more suitably at least 25 or preferably at least 50 contiguous nucleotides of said SEQ ID NO:.

In another preferred embodiment, a single exon nucleic acid probe in accordance with the third or fourth aspects of the invention is between 3kb and 25kb in length. It is preferred that said probe is no more than 3kb, suitably no more than 5kb, more suitably no more than 10kb, preferably 15kb, more preferably 20kb or, most preferably, no more than 20kb in length.

Preferably, a single exon nucleic acid probe in accordance with either the fifth or sixth aspect of the invention is DNA, preferably single-stranded DNA, RNA or PNA.

In another embodiment of either the third or fourth aspect of the invention, a single exon nucleic acid

probe is detectably labeled. Suitable detectable labels include a radionuclide, a fluorescent label or a first member of a specific binding pair. Suitable fluorescent labels include dyes such as cyanine dyes, preferably Cy3
5 and Cy5 although other suitable dyes will be known to those skilled in the art.

In a particularly preferred embodiment, a single exon nucleic acid probe in accordance with either the third or fourth aspect of the invention lacks prokaryotic and
10 bacteriophage vector sequence. In yet another embodiment, a single exon nucleic acid probe in accordance with either the third or fourth aspect of the invention lacks homopolymeric stretches of A or T.

In a fifth aspect of the invention, there is
15 provided an amplifiable nucleic acid composition, comprising:

the single exon nucleic acid probe in accordance with either of the third or fourth aspects of the invention; and at least one nucleic acid primer;

20 wherein said at least one primer is sufficient to prime enzymatic amplification of said probe.

In an sixth aspect of the invention, there is provided a method of measuring gene expression in a sample derived from human bone marrow, comprising:

25 contacting the single exon microarray in accordance with the second aspect of the invention, with a first collection of detectably labeled nucleic acids, said first collection of nucleic acids derived from mRNA of human bone marrow; and then

30 measuring the label detectably bound to each probe of said microarray.

In a seventh aspect of the invention, there is provided a method of identifying exons in a eukaryotic genome, comprising:

35 algorithmically predicting at least one exon from

genomic sequence of said eukaryote; and then
detecting specific hybridization of detectably
labeled nucleic acids to a single exon probe,

wherein said detectably labeled nucleic acids are
5 derived from mRNA from the bone marrow of said eukaryote,
said probe is a single exon probe having a fragment
identical in sequence to, or complementary in sequence to,
said predicted exon, said probe is included within a single
exon microarray in accordance with the first aspect of the
10 invention, and said fragment is selectively hybridizable at
high stringency.

In a eighth aspect of the invention, there is
provided a method of assigning exons to a single gene,
comprising:

15 identifying a plurality of exons from genomic
sequence in accordance with the seventh aspect of the
invention; and then

measuring the expression of each of said exons in
a plurality of tissues and/or cell types using
20 hybridization to single exon microarrays having a probe
with said exon,

wherein a common pattern of expression of said
exons in said plurality of tissues and/or cell types
indicates that the exons should be assigned to a single
25 gene.

In an ninth aspect of the invention, there is
provided a nucleic acid sequence as set out in any of SEQ
ID NOs: 1 - 26,012 wherein said sequence encodes a peptide.

In a tenth aspect of the invention, there is
30 provided a peptide encoded by a sequence comprising a
sequence as set out in any of SEQ ID NOs: 13,115 - 26,012,
or a complementary sequence or coding portion thereof.

In a preferred embodiment, a peptide may be
encoded by a sequence comprising a sequence set out in any
35 of SEQ ID NOS.: 1 - 13,114.

In a further aspect, the invention provides peptides comprising an amino acid sequence translated from the DNA fragments, said amino acid sequences comprising SEQ ID NOS.: 26,013 - 38,628.

5 Accordingly in a eleventh aspect of the invention there is provided a peptide comprising a sequence as set out in any of SEQ ID NOs: 26,013 - 38,628, or fragment thereof.

In another aspect, the invention provides means
10 for displaying annotated sequence, and in particular, for displaying sequence annotated according to the methods and apparatus of the present invention. Further, such display can be used as a preferred graphical user interface for electronic search, query, and analysis of such annotated
15 sequence.

Detailed Description of the Invention

20 Definitions

As used herein, the term "microarray" and phrase "nucleic acid microarray" refer to a substrate-bound collection of plural nucleic acids, hybridization to each of the plurality of bound nucleic acids being separately
25 detectable. The substrate can be solid or porous, planar or non-planar, unitary or distributed.

As so defined, the term "microarray" and phrase "nucleic acid microarray" include all the devices so called in Schena (ed.), DNA Microarrays: A Practical Approach
30 (Practical Approach Series), Oxford University Press (1999) (ISBN: 0199637768); Nature Genet. 21(1)(suppl):1 - 60 (1999); and Schena (ed.), Microarray Biochip: Tools and Technology, Eaton Publishing Company/BioTechniques Books Division (2000) (ISBN: 1881299376). As so defined, the
35 term "microarray" and phrase "nucleic acid microarray"

further include substrate-bound collections of plural nucleic acids in which the nucleic acids are distributably disposed on a plurality of beads, rather than on a unitary planar substrate, as is described, *inter alia*, in Brenner
5 *et al.*, *Proc. Natl. Acad. Sci. USA* 97(4):166501670 (2000); in such case, the term "microarray" and phrase "nucleic acid microarray" refer to the plurality of beads in aggregate.

As used herein with respect to a nucleic acid
10 microarray, the term "probe" refers to the nucleic acid that is, or is intended to be, bound to the substrate; in such context, the term "target" thus refers to nucleic acid intended to be bound thereto by Watson-Crick complementarity. As used herein with respect to solution
15 phase hybridization, the term "probe" refers to the nucleic acid of known sequence that is detectably labeled.

As used herein, the expression "probe comprising SEQ ID NO.", and variants thereof, intends a nucleic acid probe, at least a portion of which probe has either (i) the
20 sequence directly as given in the referenced SEQ ID NO., or (ii) a sequence complementary to the sequence as given in the referenced SEQ ID NO., the choice as between sequence directly as given and complement thereof dictated by the requirement that the probe hybridize to mRNA.

25 As used herein, the term "open reading frame" and the equivalent acronym "ORF" refer to that portion of an exon that can be translated in its entirety into a sequence of contiguous amino acids i.e. a nucleic acid sequence that, in at least one reading frame, does not possess stop
30 codons; the term does not require that the ORF encode the entirety of a natural protein.

As used herein, the term "amplicon" refers to a PCR product amplified from human genomic DNA, containing the predicted exon.

35 As used herein the term "exon" refers to the

consensus prediction of the various exon and gene predicting algorithms i.e. a nucleic acid sequence bioinformatically predicted to encode a portion of a natural protein.

5 As used herein, the term "peptide" refers to a sequence of amino acids. The sequences referred to as PEPTIDE SEQ ID NOS.: are the predicted peptide sequences that would be translated from one of the exons, or a portion thereof set out in exon SEQ ID NOS.: The codons
10 encoding the peptide are wholly contained within the exon.

 As used herein, a "portions" of a defined nucleotide sequence or sequences can be and, preferably, are fragments unique to that sequence or to one or a combination of those sequences. A fragment unique to a
15 nucleic acid molecule is one that is a signature for the larger nucleic acid molecule.

 As used herein, the phrase "expression of a probe" and its linguistic variants means that the ORF present within the probe, or its complement, is present
20 within a target mRNA.

 As used herein, "stringent conditions" refers to parameters well known to those skilled in the art. When a nucleic acid molecule is said to be hybridisable to another of a given sequence under "stringent conditions" it is
25 meant that it is homologous to the given sequence.

 As used herein, the phrase "specific binding pair" intends a pair of molecules that bind to one another with high specificity. Binding pairs are said to exhibit specific binding when they exhibit avidity of at least 10^7 ,
30 preferably at least 10^8 , more preferably at least 10^9 liters/mole. Nonlimiting examples of specific binding pairs are: antibody and antigen; biotin and avidin; and biotin and streptavidin.

 As used herein with respect to the visual display
35 of annotated genomic sequence, the term "rectangle" means

any geometric shape that has at least a first and a second border, wherein the first and second borders each are capable of mapping uniquely to a point of another visual object of the display.

5 As used herein, a "Mondrian" means a visual display in which a single genomic sequence is annotated with predicted and experimentally confirmed functional information.

10

Brief Description of the Drawings

The present invention is further illustrated with reference to the following non-limiting figures and
15 examples in which:

FIG. 1 illustrates a process for predicting functional regions from genomic sequence, confirming the functional activity of such regions experimentally, and associating and displaying the data so obtained in
20 meaningful and useful relationship to the original sequence data;

FIG. 2 further elaborates that portion of the process schematized in FIG. 1 for predicting functional regions from genomic sequence;

25 FIG. 3 illustrates a Mondrian visual display;

FIG. 4 presents a Mondrian showing a hypothetical annotated genomic sequence;

FIG. 5 is a histogram showing the distribution of ORF length and PCR products as obtained, with ORF length
30 shown in black and PCR product length shown in dotted lines;

FIG. 6 is a histogram showing the distribution, among exons predicted according to the methods described, of expression as measured using simultaneous two color
35 hybridization to a genome-derived single exon microarray.

The graph shows the number of sequence-verified products that were either not expressed ("0"), expressed in one or more but not all tested tissues ("1" - "9"), or expressed in all tissues tested ("10");

5 FIG. 7 is a pictorial representation of the expression of verified sequences that showed expression with signal intensity greater than 3 in at least one tissue, with: FIG. 7A showing the expression as measured by microarray hybridization in each of the 10 measured
10 tissues, and the expression as measured "bioinformatically" by query of EST, NR and SwissProt databases; with FIG. 7B showing the legend for display of physical expression (ratio) in FIG. 7A; and with FIG. 7C showing the legend for scoring EST hits as depicted in FIG. 7A;

15 FIG. 8 shows a comparison of normalized CY3 signal intensity for arrayed sequences that were identical to sequences in existing EST, NR and SwissProt databases or that were dissimilar (unknown), where black denotes the signal intensity for all sequence-verified products with a
20 BLAST Expect ("E") value of greater than $1e-30$ (1×10^{-30}) ("unknown") and a dotted line denotes sequence-verified spots with a BLAST expect ("E") value of less than $1e-30$ (1×10^{-30}) ("known");

 FIG. 9 presents a Mondrian of BAC AC008172 (bases
25 25,000 to 130,000), containing the carbamyl phosphate synthetase gene (AF154830.1); and

 FIG. 10 is a Mondrian of BAC A049839.

30 Methods and Apparatus for Predicting, Confirming,
Annotating, and Displaying Functional Regions From Genomic
Sequence Data

 FIG. 1 is a flow chart illustrating in broad
35 outline a process for predicting functional regions from

genomic sequence, confirming and characterizing the functional activity of such regions experimentally, and then associating and displaying the information so obtained in meaningful and useful relationship to the original
5 sequence data.

The initial input into process 10 of the present invention is drawn from one or more databases 100 containing genomic sequence data. Because genomic sequence is usually obtained from subgenomic fragments, the sequence
10 data typically will be stored in a series of records corresponding to these subgenomic sequenced fragments. Some fragments will have been catenated to form larger contiguous sequences ("contigs"); others will not. A finite percentage of sequence data in the database will
15 typically be erroneous, consisting *inter alia* of vector sequence, sequence created from aberrant cloning events, sequence of artificial polylinkers, and sequence that was erroneously read.

Each sequence record in database 100 will
20 minimally contain as annotation a unique sequence identifier (accession number), and will typically be annotated further to identify the date of accession, species of origin, and depositor. Because database 100 can contain nongenomic sequence, each sequence will typically
25 be annotated further to permit query for genomic sequence. Chromosomal origin, optionally with map location, can also be present. Data can be, and over time increasingly will be, further annotated with additional information, in part through use of the present invention, as described below.
30 Annotation can be present within the data records, in information external to database 100 and linked to the records thereto, or through a combination of the two.

Databases useful as genomic sequence database 100 in the present invention include GenBank, and particularly
35 include several divisions thereof, including the

htgs(draft), NT (nucleotide, command line), and NR (nonredundant) divisions. GenBank is produced by the National Institutes of Health and is maintained by the National Center for Biotechnology Information (NCBI).

5 Databases of genomic sequence from species other than human, such as mouse, rat, *Arabidopsis*, *C. elegans*, *C. briggsii*, *Drosophila*, zebra fish, and other higher eukaryotic organisms will also prove useful as genomic sequence database 100.

10 Genomic sequence obtained by query of genomic sequence database 100 is then input into one or more processes 200 for identification of regions therein that are predicted to have a biological function as specified by the user. Such functions include, but are not limited to,
15 encoding protein, regulating transcription, regulating message transport after transcription into mRNA, regulating message splicing after transcription into mRNA, of regulating message degradation after transcription into mRNA, and the like. Other functions include directing
20 somatic recombination events, contributing to chromosomal stability or movement, contributing to allelic exclusion or X chromosome inactivation, and the like.

The particular genomic sequence to be input into process 200 will depend upon the function for which
25 relevant sequence is to be identified as well as upon the approach chosen for such identification. Process step 200 can be iterated to identify different functions within a given genomic region. In such case, the input often will be different for the several iterations.

30 Sequences predicted to have the requisite function by process 200 are then input into process 300, where a subset of the input sequences suitable for experimental confirmation is identified. Experimental confirmation can involve physical and/or bioinformatic
35 assay. Where the subsequent experimental assay is

bioinformatic, rather than physical, there are fewer constraints on the sequences that can be tested, and in this latter case therefore process 300 can output the entirety of the input sequence.

5 The subset of sequences output from process 300 is then used in process 400 for experimental verification and characterization of the function predicted in process 200, which experimental verification can, and often will, include both physical and bioinformatic assay.

10 Process 500 annotates the sequence data with the functional information obtained in the physical and/or bioinformatic assays of process 400. Such annotation can be done using any technique that usefully relates the functional information to the sequence, as, for example, by
15 incorporating the functional data into the sequence data record itself, by linking records in a hierarchical or relational database, by linking to external databases, by a combination thereof, or by other means well known within the database arts. The data can even be submitted for
20 incorporation into databases maintained by others, such as GenBank, which is maintained by NCBI.

As further noted in FIG. 1, additional annotation can be input into process 500 from external sources 600.

The annotated data is then displayed in process
25 800, either before, concomitantly with, or after optional storage 700 on nontransient media, such as magnetic disk, optical disc, magnetooptical disk, flash memory, or the like.

FIG. 1 shows that the experimental data output
30 from process 400 can be used in each preceding step of process 10: e.g., facilitating identification of functional sequences in process 200, facilitating identification of an experimentally suitable subset thereof in process 300, and facilitating creation of physical and/or informational
35 substrates for, and performance of subsequent assay, of

functional sequences in process 400.

Information from each step can be passed directly to the succeeding process, or stored in permanent or interim form prior to passage to the succeeding process.

5 Often, data will be stored after each, or at least a plurality, of such process steps. Any or all process steps can be automated.

FIG. 2 further elaborates the prediction of functional sequence within genomic sequence according to
10 process 200.

Genomic sequence database 100 is first queried 20 for genomic sequence.

The sequence required to be returned by query 20 will depend, in the first instance, upon the function to be
15 identified.

For example, genomic sequences that function to encode protein can be identified *inter alia* using gene prediction approaches, comparative sequence analysis approaches, or combinations of the two. In gene prediction
20 analysis, sequence from one genome is input into process 200 where at least one, preferably a plurality, of algorithmic methods are applied to identify putative coding regions. In comparative sequence analysis, by contrast, corresponding, e.g., syntenic, sequence from a plurality of
25 sources, typically a plurality of species, is input into process 200, where at least one, possibly a plurality, of algorithmic methods are applied to compare the sequences and identify regions of least variability.

The exact content of query 20 will also depend
30 upon the database queried. For example, if the database contains both genomic and nongenomic sequence, perhaps derived from multiple species, and the function to be determined is protein coding regions in human genomic sequence, the query will accordingly require that the
35 sequence returned be genomic and derived from humans.

Query 20 can also incorporate criteria that compel return of sequence that meets operative requirements of the subsequent analytical method. Alternatively, or in addition, such operative criteria can be enforced in
5 subsequent preprocess step 24.

For example, if the function sought to be identified is protein coding, query 20 can incorporate criteria that return from genomic sequence database 100 only those sequences present within contigs sufficiently
10 long as to have obviated substantial fragmentation of any given exon among a plurality of separate sequence fragments.

Such criteria can, for example, consist of a required minimal individual genomic sequence fragment
15 length, such as 10 kb, more typically 20 kb, 30 kb, 40kb, and preferably 50 kb or more, as well as an optional further or alternative requirement that sequence from any given clone, such as a bacterial artificial chromosome ("BAC"), be presented in no more than a finite maximal
20 number of fragments, such as no more than 20 separate pieces, more typically no more than 15 fragments, even more typically no more than about 10 - 12 fragments.

Results using the present invention have shown that genomic sequence from bacterial artificial chromosomes
25 (BACs) is sufficient for gene prediction analysis according to the present invention if the sequence is at least 50 kb in length, and if additionally the sequence from any given BAC is presented in fewer than 15, and preferably fewer than 10, fragments. Accordingly, query 20 can incorporate
30 a requirement that data accessioned from BAC sequencing be in fewer than 15, preferably fewer than 10, fragments.

An additional criterion that can be incorporated into the query can be the date, or range of dates, of sequence accession. Although the process has been
35 described above as if genomic sequence database 100 were

static, it is of course understood that the genomic sequence databases need not be static, and indeed are typically updated on a frequent, even hourly, basis. Thus, as further described in Examples 1 and 2, *infra*, it is possible to query the database for newly added sequence, either newly added after an absolute date, or newly added relative to a prior analysis performed using the methods and apparatus of the present invention. In this way, the process herein described can incorporate a dynamic, temporal component.

One utility of such temporal limitation is to identify, from newly accessioned genomic sequence, the presence of novel genes, particularly those not previously identified by EST sequencing (or other sequencing efforts that are similarly based upon gene expression). As further described in Example 1, such an approach has shown that newly accessioned human genomic sequence, when analyzed for sequences that function to encode protein, readily identifies genes that are novel over those in existing EST and other expression databases. This makes the methods of the present invention extremely powerful gene discovery tools. And as would be appreciated, such gene discovery can be performed using genomic sequence from species other than human.

If query 20 incorporates multiple criteria, such as above-described, the multiple criteria can be performed as a series of separate queries or as a single query, depending in part upon the query language, the complexity of the query, and other considerations well known in the database arts.

If query 20 returns no genomic sequence meeting the query criteria, the negative result can be reported by process 22, and process 200 (and indeed, entire process 10) ended 23, as shown. Alternatively, or in addition to report and termination of the initial inquiry, a new query

20 can be generated that takes into account the initial negative result.

When query 20 returns sequence meeting the query criteria, the returned sequence is then passed to optional
5 preprocessing 24, suitable and specific for the desired analytical approach and the particular analytical methods thereof to be used in process 25.

Preprocessing 24 can include processes suitable for many approaches and methods thereof, as well as
10 processes specifically suited for the intended subsequent analysis.

Preprocessing 24 suitable for most approaches and methods will include elimination of sequence irrelevant to, or that would interfere with, the subsequent analysis.
15 Such sequence includes repetitive sequence, such as Alu repeats and LINE elements, vector sequence, artificial sequence, such as artificial polylinkers, and the like. Such removal can readily be performed by identification and subsequent masking of the undesired sequence.

Identification can be effected by comparing the
20 genomic sequence returned by query 20 with public or private databases containing known repetitive sequence, vector sequence, artificial sequence, and other artifactual sequence. Such comparison can readily be done using
25 programs well known in the art, such as CROSS_MATCH, or by proprietary sequence comparison programs the engineering of which is well within the skill in the art.

Alternatively, or in addition, undesirable, including artifactual, sequence can be identified
30 algorithmically without comparison to external databases and thereafter removed. For example, synthetic polylinker sequence can be identified by an algorithm that identifies a significantly higher than average density of known restriction sites. As another example, vector sequence can
35 be identified by algorithms that identify nucleotide or

codon usage at variance with that of the bulk of the genomic sequence.

Once identified, undesired sequence can be removed. Removal can usefully be done by masking the
5 undesired sequence as, for example, by converting the specific nucleotide references to one that is unrecognized by the subsequent bioinformatic algorithms, such as "X". Alternatively, but at present less preferred, the undesired sequence can be excised from the returned genomic sequence,
10 leaving gaps.

Preprocessing 24 can further include selection from among duplicative sequences of that one sequence of highest quality. Higher quality can be measured as a lower percentage of, fewest number of, or least densely clustered
15 occurrence of ambiguous nucleotides, defined as those nucleotides that are identified in the genomic sequence using symbols indicating ambiguity. Higher quality can also or alternatively be valued by presence in the longest contig.

20 Preprocessing 24 can, and often will, also include formatting of the data as specifically appropriate for passage to the analytical algorithms of process 25. Such formatting can and typically will include, *inter alia*, addition of a unique sequence identifier, either derived
25 from the original accession number in genomic sequence database 100, or newly applied, and can further include additional annotation. Formatting can include conversion from one to another sequence listing standard, such as conversion to or from FASTA or the like, depending upon the
30 input expected by the subsequent process.

Preprocessing, which can be optional depending upon the function desired to be identified and the informational requirements of the methods for effecting such identification, is followed by sequence processing 25,
35 where sequences with the desired function are identified

within the genomic sequence.

As mentioned above, such functions can include, but are not limited to, encoding protein, regulating transcription, regulating message transport after
5 transcription into mRNA, regulating message splicing after transcription, of regulating message degradation, and the like. Other functions include directing somatic recombination events, contributing to chromosomal stability or movement, contributing to allelic exclusion or X
10 chromosome inactivation, or the like.

The methods of the present invention are particularly useful for gene discovery, that is, for identifying, from genomic sequence, regions that function to encode genes, and in a particularly useful embodiment,
15 for identifying regions that function to encode genes not hitherto identified by expression-based or directed cloning and sequencing. In conjunction with verification using the novel single exon microarrays of the present invention, as further described below, the methods herein described
20 become powerful gene discovery tools.

Accordingly, in a preferred embodiment of the present invention, process 25 is used to identify putative coding regions. Two preferred approaches in process 25 for identifying sequence that encodes putative genes are gene
25 prediction and comparative sequence analysis.

Gene prediction can be performed using any of a number of algorithmic methods, embodied in one or more software programs, that identify open reading frames (ORFs) using a variety of heuristics, such as GRAIL, DICTION, and
30 GENEFINDER. Comparative sequence analysis similarly can be performed using any of a variety of known programs that identify regions with lower sequence variability.

As further described in Example 1, below, gene finding software programs yield a range of results. For
35 the newly accessioned human genomic sequence input in

Example 1, for example, GRAIL identified the greatest percentage of genomic sequence as putative coding region, 2% of the data analyzed; GENEFINDER was second, calling 1%; and DICTION yielded the least putative coding region, with 0.8% of genomic sequence called as coding region.

Increased reliability can be obtained when consensus is required among several such methods. Although discussed herein particularly with respect to exon calling, consensus among methods will in general increase reliability of predicting other functions as well.

Thus, as indicated by query 26, sequence processing 25, optionally with preprocessing 24, can be repeated with a different method, with consensus among such iterations determined and reported in process 27.

Process 27 compares the several outputs for a given input genomic sequence and identifies consensus among the separately reported results. The consensus itself, as well as the sequence meeting that consensus, is then stored in process 29a, displayed in process 29b, and/or output to process 300 for subsequent identification of a subset thereof suitable for assay.

Multiple levels of consensus can be calculated and reported by process 27. For example, as further described in Example 1, *infra*, process 27 can report consensus as between all specific pairs of methods of gene prediction, as consensus among any one or more of the pairs of methods of gene prediction, or as among all of the gene prediction algorithms used. Thus, in Example 1, process 27 reported that GRAIL and GENEFINDER programs agreed on 0.7% of genomic sequence, that GRAIL and DICTION agreed on 0.5% of genomic sequence, and that the three programs together agreed on 0.25% of the data analyzed. Put another way, 0.25% of the genomic sequence was identified by all three of the programs as containing putative coding region.

Furthermore, consensus can be required among

different approaches to identifying a chosen function.

For example, if the function desired to be identified is coding of protein sequence, and a first used approach to exon calling is gene prediction, the process
5 can be repeated on the same input sequence, or subset thereof, with another approach, such as comparative sequence analysis. In such a case, where comparative sequence analysis follows gene prediction, the comparison can be performed not only on genomic nucleic acid sequence,
10 but additionally or alternatively can be performed on the predicted amino acid sequence translated from the ORFs prior identified by the gene prediction approach.

Although shown as an iterative process, the multiple analyses required to achieve consensus can be done
15 in series, in parallel, or some combination thereof.

Predicted functional sequence, optionally representing a consensus among a plurality of methods and approaches for determination thereof, is passed to process
300 for identification of a subset thereof for functional
20 assay.

In the preferred embodiment of the methods of the present invention, wherein the function sought to be identified is protein coding, process 300 is used to identify a subset thereof suitable for experimental
25 verification by physical and/or bioinformatic approaches.

For example, putative ORFs identified in process 200 can be classified, or binned, bioinformatically into putative genes. This binning can be based *inter alia* upon consideration of the average number of exons/gene in the
30 species chosen for analysis, upon density of exons that have been called on the genomic sequence, and other empirical rules. Thereafter, one or more among the gene-specific ORFs can be chosen for subsequent use in gene expression assay.

35 Where such subsequent gene expression assay uses

amplified nucleic acid, considerations such as desired amplicon length, primer synthesis requirements, putative exon length, sequence GC content, existence of possible secondary structure, and the like can be used to identify
5 and select those ORFs that appear most likely successfully to amplify. Where subsequent gene expression assay relies upon nucleic acid hybridization, whether or not using amplified product, further considerations involving hybridization stringency can be applied to identify that
10 subset of sequences that will most readily permit sequence-specific discrimination at a chosen hybridization and wash stringency. One particular such consideration is avoidance of putative exons that span repetitive sequence; such sequence can hybridize spuriously to nonspecific message,
15 reducing specific signal in the hybridization.

For bioinformatic assay, there are fewer constraints on the sequences that can be tested experimentally, and in this latter case therefore process 300 can output the entirety of the input sequence.

20 The subset of sequences identified by process 300 as suitable for use in assay is then used in process 400 to create the physical and/or informational substrate for experimental verification of the predictions made in process 200, and thereafter to assay those substrates.

25 As mentioned, the methods of the present invention are particularly useful for identifying potential coding regions within genomic sequence. In a preferred embodiment of process 400, therefore, the expression of the sequences predicted to encode protein is verified. The
30 combination of the predictive and experimental methods provides a powerful gene discovery engine.

Thus, in another aspect, the present invention provides methods and apparatus for verifying the expression of putative genes identified within genomic sequence. In
35 particular, the invention provides a novel method of

verifying gene expression in which expression of predicted ORFs is measured and confirmed using a novel type of nucleic acid microarray, the genome-derived single exon nucleic acid microarrays of the present invention.

5 Putative ORFs as predicted by a consensus of gene calling, particularly gene prediction, algorithms in process 200, and as further identified as suitable by process 300, are amplified from genomic DNA using the polymerase chain reaction (PCR). Although PCR is
10 conveniently used, other amplification approaches can also be used.

Amplification schemes can be designed to capture the entirety of each predicted ORF in an amplicon with minimal additional (that is, intronic or intergenic)
15 sequence. Because ORFs predicted from human genomic sequence using the methods of the present invention differ in length, such an approach results in amplicons of varying length.

However, most predicted ORFs are shorter than 500
20 bp in length, and although amplicons of at least about 100 or 200 base pairs can be immobilized as probes on nucleic acid microarrays, early experimental results using the methods of the present invention have suggested that longer amplicons, at least about 400 or 500 base pairs, are more
25 effective. Furthermore, certain advantages derive from application to the microarray of amplicons of defined size.

Therefore, amplification schemes can alternatively, and preferably, be designed to amplify regions of defined size, preferably at least about 300, 400
30 or 500 bp, centered about each predicted ORF. Such an approach results in a population of amplicons of limited size diversity, but that typically contain intronic and/or intergenic nucleic acid in addition to putative ORF.

Conversely, somewhat fewer than 10% of ORFs
35 predicted from human genomic sequence according to the

methods of the present invention exceed 500 bp in length. Portions of such extended ORFs, preferably at least about 300,400 or 500 bp in length, can be amplified. However, it has been discovered that the percentage success at
5 amplifying pieces of such ORFs is low, and that such putative exons are more effectively amplified when larger fragments, at least about 1000 or 1500 bp, and even as large as 2000 bp are amplified.

The putative ORFs selected in process 300 are
10 thus input into one or more primer design programs, such as PRIMER3 (available online for use at <http://www-genome.wi.mit.edu/cgi-bin/primer/>), with a goal of amplifying at least about 500 base pairs of genomic sequence centered within or about ORFs predicted to be no
15 more than about 500 bp, or at least about 1000 - 1500 bp of genomic sequence for ORFs predicted to exceed 500 bp in length, and the primers synthesized by standard techniques. Primers with the requisite sequences can be purchased commercially or synthesized by standard techniques.

20 Conveniently, a first predetermined sequence can be added commonly to the ORF-specific 5' primer and a second, typically different, predetermined sequence commonly added to each 3' ORF-unique primer. This serves to immortalize the amplicon, that is, serves to permit
25 further amplification of any amplicon using a single set of primers complementary respectively to the common 5' and common 3' sequence elements. The presence of these "universal" priming sequences further facilitates later sequence verification, providing a sequence common to all
30 amplicons at which to prime sequencing reactions. The common 5' and 3' sequences further serve to add a cloning site should any of the ORFs warrant further study.

Such predetermined sequence is usefully at least about 10, 12 or 15 nt in length, and usually does not
35 exceed about 25 nt in length. The "universal" priming

sequences used in the examples presented *infra* were each 16 nt long.

The genomic DNA to be used as substrate for amplification will come from the eukaryotic species from which the genomic sequence data had originally been obtained, or a closely related species, and can conveniently be prepared by well known techniques from somatic or germline tissue or cultured cells of the organism. See, e.g., Short Protocols in Molecular Biology : A Compendium of Methods from Current Protocols in Molecular Biology, Ausubel et al. (eds.), 4th edition (April 1999), John Wiley & Sons (ISBN: 047132938X) and Maniatis et al., Molecular Cloning : A Laboratory Manual, 2nd edition (December 1989), Cold Spring Harbor Laboratory Press (ISBN: 0879693096). Many such prepared genomic DNAs are available commercially, with the human genomic DNAs additionally having certification of donor informed consent.

Although the intronic and intergenic material flanking putative coding regions in the amplicons could potentially interfere with hybridizations during microarray experiments, we have found, surprisingly, that differential expression ratios are not significantly affected. Rather, the predominant effect of exon size is to alter the absolute signal intensity, rather than its ratio. Equally surprising, the art had suggested that single exon probes would not provide sufficient signal intensity for high stringency hybridization analyses; we find that such probes not only provide adequate signal, but have substantial advantages, as herein described.

After partial purification, as by size exclusion spin column, with or without confirmation as to amplicon quality as by gel electrophoresis, each amplicon (single exon probe) is disposed in an array upon a support substrate.

Methods for creating microarrays by deposition and fixation of nucleic acids onto support substrates are well known in the art (Reviewed by Schena et al., see above).

5 Typically, the support substrate will be glass, although other materials, such as amorphous or crystalline silicon or plastics. Such plastics include polymethylacrylic, polyethylene, polypropylene, polyacrylate, polymethylmethacrylate, polyvinylchloride, 10 polytetrafluoroethylene, polystyrene, polycarbonate, polyacetal, polysulfone, celluloseacetate, cellulosenitrate, nitrocellulose, or mixtures thereof, can also be used. Typically, the support will be rectangular, although other shapes, particularly circular disks and even 15 spheres, present certain advantages. Particularly advantageous alternatives to glass slides as support substrates for array of nucleic acids are optical discs, as described in WO 98/12559.

 The amplified nucleic acids can be attached 20 covalently to a surface of the support substrate or, more typically, applied to a derivatized surface in a chaotropic agent that facilitates denaturation and adherence by presumed noncovalent interactions, or some combination thereof.

25 Robotic spotting devices useful for arraying nucleic acids on support substrates can be constructed using public domain specifications (The MGuide, version 2.0, <http://cmgm.stanford.edu/pbrown/mguide/index.html>), or can conveniently be purchased from commercial sources 30 (MicroArray GenII Spotter and MicroArray GenIII Spotter, Molecular Dynamics, Inc., Sunnyvale, CA). Spotting can also be effected by printing methods, including those using ink jet technology.

 As is well known in the art, microarrays 35 typically also contain immobilized control nucleic acids.

For controls useful in providing measurements of background signal for the genome-derived single exon microarrays of the present invention, a plurality of *E. coli* genes can readily be used. As further described in Example 1, 16 or
5 32 *E. coli* genes suffice to provide a robust measure of background noise in such microarrays.

As is well known in the art, the amplified product disposed in arrays on a support substrate to create a nucleic acid microarray can consist entirely of natural
10 nucleotides linked by phosphodiester bonds, or alternatively can include either nonnative nucleotides, alternative internucleotide linkages, or both, so long as complementary binding can be obtained in the hybridization. If enzymatic amplification is used to produce the
15 immobilized probes, the amplifying enzyme will impose certain further constraints upon the types of nucleic acid analogs that can be generated.

Although particularly described herein as using high density microarrays constructed on planar substrates,
20 the methods of the present invention for confirming the expression of ORFs predicted from genomic sequence can use any of the known types of microarrays, as herein defined, including lower density planar arrays, and microarrays on nonplanar, nonunitary, distributed substrates.

25 For example, gene expression can be confirmed using hybridization to lower density arrays, such as those constructed on membranes, such as nitrocellulose, nylon, and positively-charged derivatized nylon membranes. Further, gene expression can also be confirmed using
30 nonplanar, bead-based microarrays such as are described in Brenner *et al.*, *Proc. Natl. Acad. Sci. USA* 97(4):166501670 (2000); U.S. Patent No. 6,057,107; and U.S. Patent No. 5,736,330. In theory, a packed collection of such beads provides in aggregate a higher density of nucleic acid
35 probe than can be achieved with spotting or lithography

techniques on a single planar substrate.

Planar microarrays on solid substrates, however, provide certain useful advantages, including high throughput and compatibility with existing readers. For example, each standard microscope slide can include at least 1000, typically at least 2000, preferably 5000 and upto 10,000 - 50,000 or more nucleic acid probes of discrete sequence. The number of sequences deposited will depend on their required application.

Each putative gene can be represented in the array by a single predicted ORF. Alternatively, genes can be represented by more than one predicted ORF. For purposes of measuring differential splicing, more than one predicted ORF will be provided for a putative gene. And as is well known in the art, each probe of defined sequence, representing a single predicted ORF, can be deposited in a plurality of locations on a single microarray to provide redundancy of signal.

The genome-derived single exon microarrays described above differ in several fundamental and advantageous ways from microarrays presently used in the gene expression art, including (1) those created by deposition of mRNA-derived nucleic acids, (2) those created by *in situ* synthesis of oligonucleotide probes, and (3) those constructed from yeast genomic DNA.

Most nucleic acid microarrays that are in use for study of eukaryotic gene expression have as immobilized probes nucleic acids that are derived - either directly or indirectly - from expressed message. As discussed above, it is common, for example, for such microarrays to be derived from cDNA/EST libraries, either from those previously described in the literature, see Lennon et al., or from the *de novo* construction of "problem specific" libraries targeted at a particular biological question, R.S. Thomas et al., Cancer Res. (in press). Such

microarrays are herein collectively denominated "EST microarrays".

Such EST microarrays by definition can measure expression only of those genes found in EST libraries, shown herein to represent only a fraction of expressed genes. Furthermore, such libraries - and thus microarrays based thereupon - are biased by the tissue or cell type of message origin, by the expression levels of the respective genes within the tissues, and by the ability of the message successfully to have been reverse-transcribed and cloned.

Thus, as further discussed in Example 1, the methods of the present invention enable sequences that do not appear in EST or other expression databases to be determined - subsequently arrayed for expression measurements could not, therefore, have been represented as probes on an EST microarray. And as further demonstrated in the examples, *infra*, the remaining population of genes identified from genomic sequence by the methods of the present invention - that is, the one third of sequences that had previously been accessioned in EST or other expression databases - are biased toward genes with higher expression levels.

Representation of a message in an EST and/or cDNA library depends upon the successful reverse transcription, optionally but typically with subsequent successful cloning, of the message. This introduces substantial bias into the population of probes available for arraying in EST microarrays.

In contrast, neither reverse transcription nor cloning is required to produce the probes arrayed on the genome-derived single exon microarrays of the present invention. And although the ultimate deposition of a probe on the genome-derived single exon microarray of the present invention depends upon a successful amplification from genomic material, *a priori* knowledge of the sequence of the

desired amplicon affords greater opportunity to recover any given probe sequence recalcitrant to amplification than is afforded by the requirement for successful reverse transcription and cloning of unknown message in EST

5 approaches.

Thus, the genome-derived single exon microarrays of the present invention present a far greater diversity of probes for measuring gene expression, with far less bias, than do EST microarrays presently used in the art.

10 As a further consequence of their ultimate origin from expressed message, the probes in EST microarrays often contain poly-A (or complementary poly-T) stretches derived from the poly-A tail of mature mRNA. These homopolymeric stretches contribute to cross-hybridization, that is, to a
15 spurious signal occasioned by hybridization to the homopolymeric tail of a labeled cDNA that lacks sequence homology to the gene-specific portion of the probe.

In contrast, the probes arrayed in the genome-derived single exon microarrays of the present invention
20 lack homopolymeric stretches derived from message polyadenylation, and thus can provide more specific signal. Typically, at least about 50, 60 or 75% of the probes on the genome-derived single exon microarrays of the present invention lack homopolymeric regions consisting of A or T,
25 where a homopolymeric region is defined for purposes herein as stretches of 25, or more, typically 30 or more, identical nucleotides.

A further distinction, which also affects the specificity of hybridization, is occasioned by the typical
30 derivation of EST microarray probes from cloned material. Because much of the probe material disposed as probes on EST microarrays is excised or amplified from plasmid, phage, or phagemid vectors, EST microarrays typically include a fair amount of vector sequence, more so when the
35 probes are amplified, rather than excised, from the vector.

In contrast, the vast majority of probes in the genome-derived single exon microarrays of the present invention contain no prokaryotic or bacteriophage vector sequence, having been amplified directly or indirectly from genomic DNA. Typically, therefore, at least about 50, 60, 70 or 80% or more of individual exon-including probes disposed on a genome-derived single exon microarray of the present invention lack vector sequence, and particularly lack sequences drawn from plasmids and bacteriophage. Preferably, at least about 85, 90 or more than 90% of exon-including probes in the genome-derived single exon microarray of the present invention lack vector sequence. With attention to removal of vector sequences through preprocessing 24, percentages of vector-free exon-including probes can be as high as 95 - 99%. The substantial absence of vector sequence from the genome-derived single exon microarrays of the present invention results in greater specificity during hybridization, since spurious cross-hybridization to a probe vector sequence is reduced.

As a further consequence of excision or amplification of probes from vectors in construction of EST microarrays, the probes arrayed thereon often contain artificial sequence, derived from vector polylinker multiple cloning sites, at both 5' and 3' ends. The probes disposed upon the genome-derived single exon microarrays need have no such artificial sequence appended thereto.

As mentioned above, however, the ORF-specific primers used to amplify putative ORFs can include artificial sequences, typically 5' to the ORF-specific primer sequence, useful for "universal" (that is, independent of ORF sequence) priming of subsequent amplification or sequencing reactions. When such "universal" 5' and/or 3' priming sequences are appended to the amplification primers, the probes disposed upon the genome-derived single exon microarray will include

artificial sequence similar to that found in EST
microarrays. However, the genome-derived single exon
microarray of the present invention can be made without
such sequences, and if so constructed, presents an even
5 smaller amount of nonspecific sequence that would
contribute to nonspecific hybridization.

Yet another consequence of typical use of cloned
material as probes in EST microarrays is that such
microarrays contain probes that result from cloning
10 artifacts, such as chimeric molecules containing coding
region of two separate genes. Derived from genomic
material, typically not thereafter cloned, the probes of
the genome-derived single exon microarrays of the present
invention lack such cloning artifacts, and thus provide
15 greater specificity of signal in gene expression
measurements.

A further consequence of the cloned origin of
probes on many EST microarrays is that the individual
probes often have disparate sizes, which can cause the
20 optimal hybridization stringency to vary among probes on a
single microarray. In contrast, as discussed above, the
probes arrayed on the genome-derived single exon
microarrays of the present invention can readily be
designed to have a narrow distribution in sizes, with the
25 range of probe sizes no greater than about 10% of the
average size, typically no greater than about 5% of the
average probe size.

Because of their origin from fully- or partially-
spliced message, probes disposed upon EST arrays will often
30 include multiple exons. The percentage of such exon-
spanning probes in an EST microarray can be calculated, on
average, based upon the predicted number of exons/gene for
the given species and the average length of the immobilized
probes. For human genes, the near-complete sequence of
35 human chromosome 22, Dunham *et al.*, *Nature* 402(6761):489-95

(1999), predicts that human genes average 5.5 exons/gene. Even with probes of 200 - 500 bp, the vast majority of human EST microarray probes include more than one exon.

In contrast, by virtue of their origin from
5 algorithmically identified ORFs in genomic sequence, the probes in the genome-derived single exon microarrays of the present invention can consist of individual exons. Thus, in contrast to EST microarrays, at least about 50, 60, 70, 75, 80, 85, 95 or 99% of probes deposited in the genome-
10 derived microarray of the present invention consist of, or include, no more than one predicted ORF.

This provides the ability, not readily achieved using EST microarrays, to use the genome-derived single exon microarrays of the present invention to measure
15 tissue-specific expression of individual exons, which in turn allows differential splicing events to be detected and characterized, and in particular, allows the correlation of differential splicing to tissue-specific expression patterns.

Furthermore, the exons that are represented in
20 EST microarrays are often biased toward the 3' or 5' end of their respective genes, since sequencing strategies used for EST identification are so biased. In contrast, no such 3' or 5' bias necessarily inheres in the selection of exons
25 for disposition on the genome-derived single exon microarrays of the present invention.

Conversely, the probes provided on the genome-derived single exon microarrays of the present invention typically, but need not necessarily, include intronic
30 and/or intergenic sequence that is absent from EST microarrays, which are derived from mature mRNA. Typically, at least about 50, 60, 70, 80 or 90% of the exon-including probes on the genome-derived single exon microarrays of the present invention include sequence drawn
35 from noncoding regions. As discussed above, the additional

presence of noncoding region does not significantly interfere with measurement of gene expression, and provides the additional opportunity to assay prespliced RNA, and thus measure such phenomena such as nuclear export control.

5 The genome-derived single exon microarrays of the present invention are also quite different from *in situ* synthesis microarrays, where probe size is severely constrained by inadequacies in the photolithographic synthesis process.

10 Typically, probes arrayed on *in situ* synthesis microarrays are limited to a maximum of about 25 bp. As a well known consequence, hybridization to such chips must be performed at low stringency. In order, therefore, to achieve unambiguous sequence-specific hybridization
15 results, the *in situ* synthesis microarray requires substantial redundancy, with concomitant programmed arraying for each probe of probe analogues with altered (*i.e.*, mismatched) sequence.

 In contrast, the longer probe length of the
20 genome-derived single exon microarrays of the present invention allows much higher stringency hybridization and wash. Typically, therefore, exon-including probes on the genome-derived single exon microarrays of the present invention average at least about 100, 200, 300, 400 or
25 500 bp in length. By obviating the need for substantial probe redundancy, this approach permits a higher density of probes for discrete exons or genes to be arrayed on the microarrays of the present invention than can be achieved for *in situ* synthesis microarrays.

30 A further distinction is that the probes in *in situ* synthesis microarrays typically are covalently linked to the substrate surface. In contrast, the probes disposed on the genome-derived microarray of the present invention typically are, but need not necessarily be, bound
35 noncovalently to the substrate.

Furthermore, the short probe size on *in situ* microarrays causes large percentage differences in the melting temperature of probes hybridized to their complementary target sequence, and thus causes large percentage differences in the theoretically optimum stringency across the array as a whole.

In contrast, the larger probe size in the microarrays of the present invention create lower percentage differences in melting temperature across the range of arrayed probes.

A further significant advantage of the microarrays of the present invention over *in situ* synthesized arrays is that the quality of each individual probe can be confirmed before deposition. In contrast, the quality of probes cannot be assessed on a probe-by-probe basis for the *in situ* synthesized microarrays presently being used.

The genome-derived single exon microarrays of the present invention are also distinguished over, and present substantial benefits over, the genome-derived microarrays from lower eukaryotes such as yeast. Lashkari et al., *Proc. Natl. Acad. Sci. USA* 94:13057-13062 (1997).

Only about 220 - 250 of the 6100 or so nuclear genes in *Saccharomyces cerevisiae* - that is, only about 4 - 5% - have standard, spliceosomal, introns, Lopez et al., *Nucl. Acids Res.* 28:85-86 (2000); Spingola et al., *RNA* 5(2):221-34 (1999). Furthermore, the entire yeast genome has already been sequenced. These two facts permit the ready amplification and disposition of single-ORF amplicons on such microarray without the requirement for antecedent use of gene prediction and/or comparative sequence analyses.

Thus, a significant aspect of the present invention is the ability to identify and to confirm expression of predicted coding regions in genomic sequence

drawn from eukaryotic organisms that have a higher percentage of genes having introns than do yeast such as *Saccharomyces cerevisiae*, particularly in genomic sequence drawn from eukaryotes in which at least about 10, 20 or 50% of protein-encoding genes have introns. In preferred embodiments, the methods and apparatus of the present invention are used to identify and confirm expression of novel genes from genomic sequence of eukaryotes in which the average number of introns per gene is at least about one, two or three or more.

After the physical substrate is prepared, experimental verification of predicted function is performed.

In a preferred embodiment of the present invention, where the function sought to be identified in genomic sequence is protein coding, experimental verification is performed by measuring expression of the putative ORFs, typically through nucleic acid hybridization experiments, and in particularly preferred embodiments, through hybridization to genome-derived single exon microarrays prepared as above- described.

Expression is conveniently measured and expressed for each probe in the microarray as a ratio of the expression measured concurrently in a plurality of mRNA sources, according to techniques well known in the microarray art, Reviewed in Schena et al., and as further described in Example 2, below. The mRNA source for the reference against which specific expression is measured can be drawn from a homogeneous mRNA source, such as a single cultured cell-type, or alternatively can be heterogeneous, as from a pool of mRNA derived from multiple tissues and/or cell types, as further described in Example 2, *infra*.

mRNA can be prepared by standard techniques, see Ausubel et al. and Maniatis et al., or purchased commercially. The mRNA is then typically reverse-

transcribed in the presence of labeled nucleotides: the index source (that in which expression is desired to be measured) is reverse transcribed in the presence of nucleotides labeled with a first label, typically a fluorophore (fluorochrome; fluor; fluorescent dye); the reference source is reverse transcribed in the presence of a second label, typically a fluorophore, typically fluorometrically-distinguishable from the first label. As further described in Example 2, *infra*, Cy3 and Cy5 dyes prove particularly useful in these methods. After partial purification of the index and reference targets, hybridization to the probe array is conducted according to standard techniques, typically under a coverslip.

After wash, microarrays are conveniently scanned using a commercial microarray scanning device, such as a Gen3 Scanner (Molecular Dynamics, Sunnyvale, CA). Data on expression is then passed, with or without interim storage, to process 500, where the results for each probe are related to the original sequence.

Often, hybridization of target material to the genome-derived single exon microarray will identify certain of the probes thereon as of particular interest. Thus, it is often desirable that the user be able readily to obtain sufficient quantities of an individual probe, either for subsequent arrayed deposition upon an additional support substrate, often as part of a microarray having a plurality of probes so identified, or alternatively or additionally as a solitary solid-phase or solution-phase probe, for further use.

Thus, in another aspect, the present invention provides compositions and kits for the ready production of nucleic acids identical in sequence to, or substantially identical in sequence to, probes on the genome-derived single exon microarrays of the present invention.

In this aspect, a small quantity of each probe is

disposed, typically without attachment to substrate, in a spatially-addressable ordered set, typically one per well of a microtiter dish. Although a 96 well microtiter plate can be used, greater efficiency is obtained using higher density arrays, such as are provided by microtiter plates having 384, 864, 1536, 3456, 6144, or 9600 wells, and although microtiter plates having physical depressions (wells) are conveniently used, any device that permits addressable withdrawal of reagent from fluidly-noncommunicating areas can be used.

In this aspect of the invention, therefore, a fluidly noncommunicating addressable ordered set of individual probes, corresponding to those on a genome-derived single exon microarray, is provided, with each probe in sufficient quantity to permit amplification, such as by PCR. As earlier mentioned, the ORF-specific 5' primers used for genomic amplification can have a first common sequence added thereto, and the ORF-specific 3' primers used for genomic amplification can have a second, different, common sequence added thereto, thus permitting, in this preferred embodiment, the use of a single set of 5' and 3' primers to amplify any one of the probes from the amplifiable ordered set.

Each discrete amplifiable probe can also be packaged with amplification primers, solutes, buffers, etc., and can be provided in dry (e.g., lyophilized) form or wet, in the latter case typically with addition of agents that retard evaporation.

In another aspect of the present invention, a genome-derived single-exon microarray is packaged together with such an ordered set of amplifiable probes corresponding to the probes, or one or more subsets of probes, thereon. In alternative embodiments, the ordered set of amplifiable probes is packaged separately from the genome-derived single exon microarray.

In some embodiments, the microarray and/or ordered probe set are further packaged with recordable media that provide probe identification and addressing information, and that can additionally contain annotation
5 information, such as gene expression data. Such recordable media can be packaged with the microarray, with the ordered probe set, or with both.

If the microarray is constructed on a substrate that incorporates recordable media, such as is described in
10 international patent application no. WO 98/12559, then separate packaging of the genome-derived single exon microarray and the bioinformatic information is not required.

The amount of amplifiable probe material should
15 be sufficient to permit at least one amplification sufficient for subsequent hybridization assay.

Although the use of high density genome-derived microarrays on solid planar substrates is presently a preferred approach for the physical confirmation and
20 characterization of the expression of sequences predicted to encode protein, other types of microarrays (as herein defined) can also be used.

Furthermore, as earlier mentioned, experimental verification of the function predicted from genomic
25 sequence in process 200 can be bioinformatic, rather than, or additional to, physical verification.

For example, where the function desired to be identified is protein coding, the predicted ORFs can be compared bioinformatically to sequences known or suspected
30 of being expressed.

Thus, the sequences output from process 300 (or process 200), can be used to query expression databases, such as EST databases, SNP ("single nucleotide polymorphism") databases, known cDNA and mRNA sequences,
35 SAGE ("serial analysis of gene expression") databases, and

more generalized sequence databases that allow query for expressed sequences. Such query can be done by any sequence query algorithm, such as BLAST ("basic local alignment search tool"). The results of such query -
5 including information on identical sequences and information on nonidentical sequences that have diffuse or focal regions of sequence homology to the query sequence - can then be passed directly to process 500, or used to inform analyses subsequently undertaken in process 200,
10 process 300, or process 400.

Experimental data, whether obtained by physical or bioinformatic assay in process 400, is passed to process 500 where it is usefully related to the sequence data itself, a process colloquially termed "annotation". Such
15 annotation can be done using any technique that usefully relates the functional information to the sequence, as, for example, by incorporating the functional data into the record itself, by linking records in a hierarchical or relational database, by linking to external databases, or
20 by a combination thereof. Such database techniques are well within the skill in the art.

The annotated sequence data can be stored locally, uploaded to genomic sequence database 100, and/or displayed 800.

25 The methods and apparatus of the present invention rapidly produce functional information from genomic sequence. Coupled with the escalating pace at which sequence now accumulates, the rapid pace of sequence annotation produces a need for methods of displaying the
30 information in meaningful ways.

FIG. 3 shows visual display 80 presenting a single genomic sequence annotated according to the present invention. Because of its nominal resemblance to artistic works of Piet Mondrian, visual display 80 is alternatively
35 described herein as a "Mondrian".

Each of the visual elements of display 80 is aligned with respect to the genomic sequence being annotated (hereinafter, the "annotated sequence"). Given the number of nucleotides typically represented in an annotated sequence, representation of individual nucleotides would rarely be readable in hard copy output of display 80. Typically, therefore, the annotated sequence is schematized as rectangle 89, extending from the left border of display 80 to its right border. By convention herein, the left border of rectangle 89 represents the first nucleotide of the sequence and the right border of rectangle 89 represents the last nucleotide of the sequence.

As further discussed below, however, the Mondrian visual display of annotated sequence can serve as a convenient graphical user interface for computerized representation, analysis, and query of information stored electronically. For such use, the individual nucleotides can conveniently be linked to the X axis coordinate of rectangle 89. This permits the annotated sequence at any point within rectangle 89 readily to be viewed, either automatically - for example, by time-delayed appearance of a small overlaid window upon movement of a cursor or other pointer over rectangle 89 - or through user intervention, as by clicking a mouse or other pointing device at a point in rectangle 89.

Visual display 80 is generated after user specification of the genomic sequence to be displayed. Such specification can consist of or include an accession number for a single clone (e.g., a single BAC accessioned into GenBank), wherein the starting and stopping nucleotides are thus absolutely identified, or alternatively can consist of or include an anchor or fulcrum point about which a chosen range of sequence is anchored, thus providing relative endpoints for the

sequence to be displayed. For example, the user can anchor such a range about a given chromosomal map location, gene name, or even a sequence returned by query for similarity or identity to an input query sequence. When visual
5 display 80 is used as a graphical user interface to computerized data, additional control over the first and last displayed nucleotide will typically be dynamically selectable, as by use of standard zooming and/or selection tools.

10 Field 81 of visual display 80 is used to present the output from process 200, that is, to present the bioinformatic prediction of those sequences having the desired function within the genomic sequence. Functional sequences are typically indicated by at least one rectangle
15 83 (83a, 83b, 83c), the left and right borders of which respectively indicate, by their X-axis coordinates, the starting and ending nucleotides of the region predicted to have function.

Where a single bioinformatic method or approach
20 identifies a plurality of regions having the desired function, a plurality of rectangles 83 is disposed horizontally in field 81. Where multiple methods and/or approaches are used to identify function, each such method and/or approach can be represented by its own series of
25 horizontally disposed rectangles 83, each such horizontally disposed series of rectangles offset vertically from those representing the results of the other methods and approaches.

Thus, rectangles 83a in FIG. 3 represent the
30 functional predictions of a first method of a first approach for predicting function, rectangles 83b represent the functional predictions of a second method and/or second approach for predicting that function, and rectangles 83c represent the predictions of a third method and/or
35 approach.

Where the function desired to be identified is protein coding, field 81 is used to present the bioinformatic prediction of sequences encoding protein. For example, rectangles 83a can represent the results from
5 GRAIL or GRAIL II, rectangles 83b can represent the results from GENEFINDER, and rectangles 83c can represent the results from DICTION.

Optionally, and preferably, rectangles 83 collectively representing predictions of a single method
10 and/or approach are identically colored and/or textured, and are distinguishable from the color and/or texture used for a different method and/or approach.

Alternatively, or in addition, the color, hue, density, or texture of rectangles 83 can be used further to
15 report a measure of the bioinformatic reliability of the prediction. For example, many gene prediction programs will report a measure of the reliability of prediction. Thus, increasing degrees of such reliability can be indicated, e.g., by increasing density of shading. Where
20 display 80 is used as a graphical user interface, such measures of reliability, and indeed all other results output by the program, can additionally or alternatively be made accessible through linkage from individual rectangles 83, as by time-delayed window ("tool tip" window), or by
25 pointer (e.g., mouse)-activated link.

As earlier described, increased predictive reliability can be achieved by requiring consensus among methods and/or approaches to determining function. Thus, field 81 can include a horizontal series of rectangles 83
30 that indicate one or more degrees of consensus in predictions of function.

Although FIG. 3 shows three series of horizontally disposed rectangles in field 81, display 80 can include as few as one such series of rectangles and as
35 many as can discriminably be displayed, depending upon the

number of methods and/or approaches used to predict a given function.

Furthermore, field 81 can be used to show predictions of a plurality of different functions.

5 However, the increased visual complexity occasioned by such display makes more useful the ability of the user to select a single function for display. When display 80 is used as a graphical user interface for computer query and analysis, such function can usefully be indicated and user-
10 selectable, as by a series of graphical buttons or tabs (not shown in FIG. 3).

Rectangle 89 is shown in FIG. 3 as including interposed rectangle 84. Rectangle 84 represents the portion of annotated sequence for which predicted
15 functional information has been assayed physically, with the starting and ending nucleotides of the assayed material indicated by the X axis coordinates of the left and right borders of rectangle 84. Rectangle 85, with optional inclusive circles 86 (86a, 86b, and 86c) displays the
20 results of such physical assay.

Although a single rectangle 84 is shown in FIG. 3, physical assay is not limited to just one region of annotated genomic sequence. It is expected that an increasing percentage of regions predicted to have function
25 by process 200 will be assayed physically, and that display 80 will accordingly, for any given genomic sequence, have an increasing number of rectangles 84 and 85, representing an increased density of sequence annotation.

Where the function desired to be identified is
30 protein coding, rectangle 84 identifies the sequence of the probe used to measure expression. In embodiments of the present invention where expression is measured using genome-derived single exon microarrays, rectangle 84 identifies the sequence included within the probe
35 immobilized on the support surface of the microarray. As

noted *supra*, such probe will often include a small amount of additional, synthetic, material incorporated during amplification and designed to permit reamplification of the probe, which sequence is typically not shown in display 80.

5 Rectangle 87 is used to present the results of bioinformatic assay of the genomic sequence. For example, where the function desired to be identified is protein coding, process 400 can include bioinformatic query of expression databases with the sequences predicted in
10 process 200 to encode exons. And as earlier discussed, because bioinformatic assay presents fewer constraints than does physical assay, often the entire output of process 200 can be used for such assay, without further subsetting thereof by process 300. Therefore, rectangle 87 typically
15 need not have separate indicators therein of regions submitted for bioinformatic assay; that is, rectangle 87 typically need not have regions therein analogous to rectangles 84 within rectangle 89.

 Rectangle 87 as shown in FIG. 3 includes smaller
20 rectangles 880 and 88. Rectangles 880 indicate regions that returned a positive result in the bioinformatic assay, with rectangles 88 representing regions that did not return such positive results. Where the function desired to be predicted and displayed is protein coding, rectangles 880
25 indicate regions of the predicted exons that identify sequence with significant similarity in expression databases, such as EST, SNP, SAGE databases, with rectangles 88 indicating genes novel over those identified in existing expression data bases.

30 Rectangles 880 can further indicate, through color, shading, texture, or the like, additional information obtained from bioinformatic assay.

 For example, where the function assayed and displayed is protein coding, the degree of shading of
35 rectangles 880 can be used to represent the degree of

sequence similarity found upon query of expression databases. The number of levels of discrimination can be as few as two (identity, and similarity, where similarity has a user-selectable lower threshold). Alternatively, as
5 many different levels of discrimination can be indicated as can visually be discriminated.

Where display 80 is used as a graphical user interface, rectangles 880 can additionally provide links directly to the sequences identified by the query of
10 expression databases, and/or statistical summaries thereof. As with each of the precedingly-discussed uses of display 80 as a graphical user interface, it should be understood that the information accessed via display 80 need not be resident on the computer presenting such display, which
15 often will be serving as a client, with the linked information resident on one or more remotely located servers.

Rectangle 85 displays the results of physical assay of the sequence delimited by its left and right
20 borders.

Rectangle 85 can consist of a single rectangle, thus indicating a single assay, or alternatively, and increasingly typically, will consist of a series of rectangles (85a, 85b, 85c) indicating separate physical
25 assays of the same sequence.

Where the function assayed is gene expression, and where gene expression is assayed as herein described using simultaneous two-color fluorescent detection of hybridization to genome-derived single exon microarrays,
30 individual rectangles 85 can be colored to indicate the degree of expression relative to control. Conveniently, shades of green can be used to depict expression in the sample over control values, and shades of red used to depict expression less than control, corresponding to the
35 spectra of the Cy3 and Cy5 dyes conventionally used for

respective labeling thereof. Additional functional information can be provided in the form of circles 86 (86a, 86b, 86c), where the diameter of the circle can be used to indicate expression intensity. As discussed *infra*, such
5 relative expression (expression ratios) and absolute expression (signal intensity) can be expressed using normalized values.

Where display 80 is used as a graphical user interface, rectangle 85 can be used as a link to further
10 information about the assay. For example, where the assay is one for gene expression, each rectangle 85 can be used to link to information about the source of the hybridized mRNA, the identity of the control, raw or processed data from the microarray scan, or the like.

15 FIG. 4 is rendition of display 80 representing gene prediction and gene expression for a hypothetical BAC, showing conventions used in the Examples presented *infra*. BAC sequence ("Chip seq.") 89 is presented, with the physically assayed region thereof (corresponding to
20 rectangle 84 in FIG. 3) shown in white. Algorithmic gene predictions are shown in field 81, with predictions by GRAIL shown, predictions by GENEFINDER, and predictions by DICTION shown. Within rectangle 87, regions of sequence that, when used to query expression databases, return
25 identical or similar sequences ("EST hit") are shown as white rectangles (corresponding to rectangles 880 in FIG. 3), gray indicates low homology, and black indicates unknowns (where black and gray would correspond to rectangles 88 in FIG. 3).

30 Although FIGS. 3 and 4 show a single stretch of sequence, uninterrupted from left to right, longer sequences are usefully represented by vertical stacking of such individual Mondrians, as shown in FIGS. 9 and 10.

35 Single Exon Probes Useful For Measuring Gene Expression

The methods and apparatus of the present invention rapidly produce functional information from genomic sequence. Where the function to be identified is protein coding, the methods and apparatus of the present invention rapidly identify and confirm the expression of portions of genomic sequence that function to encode protein. As a direct result, the methods and apparatus of the present invention rapidly yield large numbers of single-exon nucleic acid probes, the majority from previously unknown genes, each of which is useful for measuring and/or surveying expression of a specific gene in one or more tissues or cell types.

It is, therefore, another aspect of the present invention to provide genome-derived single exon nucleic acid probes useful for gene expression analysis, and particularly for gene expression analysis by microarray.

Using the methods and genome-derived single-exon microarrays of the present invention, we have for example readily identified a large number of unique ORFs from human genomic sequence. Using single exon probes that encompass these ORFs, we have demonstrated, through microarray hybridization analysis, the expression of 13,114 of these ORFs in bone marrow.

As would immediately be appreciated by one of skill in the art, each single exon probe having demonstrable expression in bone marrow is currently available for use in measuring the level of its ORF's expression in bone marrow.

Because bone marrow is the tissue in which blood cells originate, diseases of the bone marrow are a significant cause of human morbidity and mortality. Increasingly, genetic factors are being found that contribute to predisposition, onset, and/or aggressiveness of most, if not all, of these diseases. Although mutations

in single genes have in some cases been identified as causal - notably in the thalassemias and sickle cell anemia - disorders of the bone marrow are, for the most part, believed to have polygenic etiologies.

5 For example, cancers that originate in the bone marrow and lymphatic tissues such as the lymphomas, leukemias, and myeloma have been recognized as a major health concern. An estimated 632,000 Americans are presently living with lymphoma, leukemia or myeloma, and
10 over 110,000 new cases are anticipated each year. The new cases alone account for 11% of all cancer cases reported in the United States.

 Lymphoma is a general term for a group of cancers of lymphocytes that manifest in the tissues of the
15 lymphatic system. Eventually, monoclonal proliferation crowds out healthy cells and creates tumors which enlarge lymph nodes. Approximately 450,000 members of the U.S. population are living with lymphoma: 160,000 with Hodgkin disease (HD) and 290,000 with non-Hodgkin lymphoma.

20 Hodgkin disease (HD) is a specialized form of lymphoma, and represent about 8% of all lymphomas. HD can be distinguish in tissues by the presence of an abnormal cell called the Reed-Sternberg cell. Incidence rates of HD are higher in adolescents and young adults, but HD is
25 considered to be one of the most curable forms of cancer. Symptoms of HD include painless swelling of lymph glands, fatigue, recurrent high fever, sweating at night, skin irritations and loss of weight.

 Although an infectious etiology has been proposed
30 to account for the disproportionate incidence of HD among siblings reared together - particularly an association with Epstein Barr Virus (EBV) - multiple genetic contributions have also been suggested.

 As early as 1986, linkage to HLA was suggested,
35 with Klitz et al., Am. J. Hum. Genet. 54: 497-505 (1994)

reporting an overall association of the nodular sclerosing (NSHD) group with the HLA class II region. Results of the study suggested that susceptibility to NSHD is influenced by more than 1 locus within the class II region. Through a literature search, Shugart and Collins (2000), *Europ. J. Hum. Genet.* 8: 460-463 (2000), performed a combined segregation and linkage analysis on 59 nuclear families with HD and concluded that HD is most likely determined by both an HLA-associated major gene and other non-HLA genetic factors, in conjunction with environmental effects.

Non-Hodgkin lymphoma (NHL) is a malignant monoclonal proliferation of the lymphoid cells in the immune system, including bone marrow, spleen, liver and GI tract. The pathologic classification of NHL continues to evolve, reflecting new insights into the cells of origin and the biologic bases of these heterogeneous diseases. The course of NHL varies from indolent and initially well tolerated to rapidly fatal. Furthermore, common clinical symptoms of NHL, but rare in HD, are congestion and edema of the face and neck and ureteral compression.

Non-Hodgkin lymphoma (NHL) has been linked to a variety of specific genetic defects, including 26 mutated genes and at least 9 identified chromosomal translocations. Among the mutated genes are: ALK (2p23); API2 (MIHC, cIAP2) (11q22-q23); API4 (survivin, SVV) (17q25(?)); ATM (ATA, ATC) (11q22.3); BCL1 (11q13.3); BCL10 (CLAP, CIPER) (1p22); BCL2 (18q21.3); BCL6 (LAZ3, ZNF51) (3q27); BLYM (1p32); BMI1 (10p13); CCND1 (D11S287E, Cyclin D, PRAD1) (11q13); CD44 (MDU3, HA, MDU2) (11pter-p13); FRAT1 (10q23-q24(?)); FRAT2 (GBP) (10(?)); IL6 (IFNB2) (7p21); IRF4 (MUM1, LSIRF) (6p25-p23); LCP1 (PLS2) (13q14.1-q14.3); MALT1 (MLT) (18q21); MUC1 (PUM, PEM) (1q21); MYBL1 (AMYB, A-MYB) (8q22); MYC (CMYC, C-MYC) (8q24.12-q24.13); NBS1 (8q21); NPM1 (B23) (5q35); PCNA (20p12); TIAM1 (21q22.1); and TP53 (p53, P53) (17q13.1).

Among the chromosomal abnormalities are: t(1;14)

(p22;q32); t(14;18)(q32;q21); t(3;14)(q27;q32);
t(6;14)(p25;q32); t(11;18)(q21;q21); t(1;14)(q21;q32);
t(2;5)(p23;q35); add(14q32) / dup(14p32); and
t(11;14)(q13;q32).

5 Additional genetic loci, as yet undiscovered, are
believed to account for other occurrences of NHL.

As another example, acute leukemia is a malignant
disease of blood-forming tissues such as the bone marrow.
It is characterized by the uncontrolled growth of white
10 blood cells. As a result, immature myeloid cells (in acute
myelogenous leukemia (AML)) or lymphoid cells (in acute
lymphocytic leukemia (ALL)) rapidly accumulate and
progressively replace the bone marrow; diminished
production of normal red cells, white cells, and platelets
15 ensues. This loss of normal marrow function in turn gives
rise to the typical clinical complications of leukemia:
anemia, infection, and bleeding.

If untreated, ALL is rapidly fatal; most patients
die within several months of diagnosis. With appropriate
20 therapy, many patients can be cured. The survival rate for
patients diagnosed with AML or ALL is 14% and 58%
respectively. However, the incidences of AML is expected
to be greater than ALL: an estimated 10,000 new cases of
AML, predominantly in older adults, is anticipated in the
25 U.S. alone, whereas 3,100 new cases of ALL are expected,
with 1,500 of these new cases occurring among children.

The etiology of acute leukemia is not known.
Although human T-cell lymphotropic virus type I (HTLV-I), a
causative agent of adult T-cell leukemia, and HTLV-II,
30 obtained from several patients with a syndrome resembling
hairy cell leukemia, have been isolated, the etiologic link
between HTLV and malignancy is uncertain. There is,
however, evidence which suggests a genetic predisposition
to incidences of acute leukemia.

35 For example, genetic disorders such as Fanconi

anemia and Down syndrome appear to increase risk of acute leukemia, specifically, AML. Evidence supporting a chromosome 21 locus for acute myelogenous leukemia (AML) includes the finding of linkage to 21q22.1-q22.2 in a
5 family with a platelet disorder and propensity to develop AML (Ho et al., Blood 87: 5218-5224 (1996), an increased incidence of leukemia in Down syndrome, and frequent somatic translocation in leukemia involving the CBFA gene on 21q22.3. In addition, Horwitz et al., Am. J. Hum.
10 Genet. 61:873-881 (1997), suggest that a gene on 16q22 may be a second cause of acute myelogenous leukemia. Nonparametric linkage analysis gave a P-value of 0.00098 for the conditional probability of linkage. Mutational analysis excluded expansion of the AT-rich minisatellite
15 repeat FRA16B fragile site and the CAG trinucleotide repeat in the E2F-4 transcription factor. Large CAG repeat expansion was excluded as a cause of leukemia in this family.

Similarly, acute lymphoblastic leukemia (ALL) has
20 been suggested to have a genetic predisposition. In particular, linkage to chromosome 9p has been reported by a number of groups. Chilcote et al., New Eng. J. Med. 313: 286-291 (1985), found that 6 of 8 patients with clinical features of lymphomatous ALL (LALL), a distinct category of
25 ALL of T-cell lineage, had karyotypic abnormalities leading to loss of bands 9p22-p21. The mechanisms varied and included deletions, unbalanced translocations, and loss of the entire chromosome; only 1 of 57 patients without LALL had an abnormality of chromosome 9 at diagnosis. Kowalczyk
30 et al., Cancer Genet. Cytogenet. 9:383-385 (1981), had earlier found changes in 9p in a subgroup of ALL cases. Chilcote et al. (1985) pointed out that there is a fragile site at 9p21 and raised the question of familial predisposition on this basis. This fragile site is the
35 breakpoint in the translocation t(9;11)(p21-22;q23), which

is associated with acute nonlymphocytic leukemia with monocytic features, ANLL-AMoL-M5a. In a large series, Murphy et al., New Eng. J. Med. 313:1611 (1985), confirmed an abnormality of 9p in 10 to 11% of cases (33 out of more than 300) of acute lymphoblastic leukemia. The breakpoints in 9p clustered in the p22-p21 region. They could not, however, corroborate the specific association with T-cell origin or so-called lymphomatous clinical features. In addition, Taki et al., Proc. Natl. Acad. Sci. USA 96:14535 (1999), recently identified AF5q31, a new AF4-related gene, fused to MLL in infant ALL with ins(5;11)(q31;q13q23), and suspects that AF5q31 and AF4 might define a new family particularly involved in the pathogenesis of 11q23-associated-ALL.

As yet a further example of a disease affecting bone marrow with likely polygenic etiology is multiple myeloma (MM).

MM is a cancer of plasma cells, the final differentiated stage of B lymphocyte maturation. The malignant clone proliferates in the bone marrow and frequently invades the adjacent bone, producing extensive skeletal destruction that results in bone pain and fractures. Anemia, hypercalcemia, and renal failure are some clinical manifestations associated with MM.

MM causes 1% of all cancer deaths in Western countries. A genetic component to its etiology is suggested by disparate incidence among various groups in the country. Its incidence is higher in men than in women, in people of African descent relative to the U.S. population at large, and in older adults as compared to the young. It has been estimated that 14,000 new cases of myeloma will be diagnosed in the U.S., and over 11,000 persons will die from MM within the year.

Although, Kaposi's sarcoma-associated herpes virus has been associated with MM (Retig et al., Science

276:1851 (1997)), there is evidence that chromosomal abnormalities, such as the deletion of 13q14 and rearrangements of 14q increase the proliferation of myeloma cells.

5 Up to 30% of patients who suffer with MM have a balanced translocation, t(4;14)(p16.3;q32), that places the fibroblast growth factor receptor 3 (FGFR3) gene under the control of IgH promoter elements (Chesi et al., Nat. Genet. 16:260 (1997)). This results in increased expression of
10 FGFR3, a member of a family of tyrosine kinase receptors implicated in control of cellular proliferation.

According to Zoger et al., Blood 95:1925 (2000), monoallelic deletions of the retinoblastoma-1 (rb-1) gene and the D13S319 locus were observed in 48 of 104 patients
15 (46.2%) and in 28 of 72 (38.9%) patients, respectively, with newly diagnosed MM. Fluorescence in situ hybridization (FISH) studies found that 13q14 was deleted in all 17 patients with karyotypic evidence of monosomy 13 or deletion of 13q but also in 9 of 19 patients with
20 apparently normal karyotypes. Patients with a 13q14 deletion were more likely to have higher serum levels of beta(2)-microglobulin (P=0.059) and a higher percentage of bone marrow plasma cells (P=0.085) than patients with a normal 13q14 status on FISH analysis. In patients with a
25 deletion of 13q14, myeloma cell proliferation was markedly increased. The presence of a 13q14 deletion on FISH analysis was associated with a significantly lower rate of response to conventional-dose chemotherapy (40.8% compared with 78.6%; P =.009) and a shorter overall
30 survival (24.2 months compared with > 60 months; P <.005) than in patients without the deletion.

There are numerous other mutated genes and chromosomal abnormalities that may predispose to MM. Examples of such genes are: B2M (15q21-q22); CCND1
35 (D11S287E, Cyclin D, PRAD1)(11q13); CD19 (16p11.2); HGF

(HPTA) (7q21.1); IL6 (IFNB2) (7p21); IRF4 (MUM1, LSIRF) (6p25-p23); LTA (TNFB, LT) (6p21.3); SDC1 (2p24.1); and TNF (TNFA, TNFSF2, DIF) (6p21.3). Examples of chromosomal abnormalities include: t(6;14) (p25;q32) and
5 t(11;14) (q13;q32).

Other significant diseases or disorders of the bone marrow are also believed, or likely to have, a genetic, typically polygenic, etiologic component. These diseases include, for example, chronic myeloid leukemia,
10 chronic lymphoid leukemia, polycythemia vera, myelofibrosis, primary thrombocythemia, myelodysplastic syndromes, Wiskott-Aldrich, lymphoproliferative syndrome, aplastic anemia, Fanconi anemia, Down syndrome, sickle cell disease, thalassemia, granulocyte disorders, Kostmann
15 syndrome, chronic granulomatous disease, Chediak-Higashi syndrome, platelet disorders, Glanzmann thrombasthenia, Bernard-Soulier syndrome, metabolic storage diseases, osteoporosis, congenital hemophagocytic syndrome.

The human genome-derived single exon nucleic acid
20 probes and microarrays of the present invention are useful for predicting, diagnosing, grading, staging, monitoring and prognosing diseases of human bone marrow, particularly those diseases with polygenic etiology. With each of the single exon probes described herein shown to be expressed
25 at detectable levels in human bone marrow, and with about 2/3 of the probes identifying novel genes, the single exon microarrays of the present invention provide exceptionally high informational content for such studies.

For example, diagnosis, grading, and/or staging
30 of a disease can be based upon the quantitative relatedness of a patient gene expression profile to one or more reference expression profiles known to be characteristic of a given bone marrow disease, or to specific grades or stages thereof.

35 In one embodiment, the patient gene expression

profile is generated by hybridizing nucleic acids obtained directly or indirectly from transcripts expressed in the patient's bone marrow (or cells cultured therefrom) to the genome-derived single exon microarray of the present
5 invention. Reference profiles are obtained similarly by hybridizing nucleic acids obtained directly or indirectly from transcripts expressed in the bone marrow of individuals with known disease. Methods for quantitatively relating gene expression profiles, without regard to the
10 function of the protein encoded by the gene, are disclosed in WO 99/58720, incorporated herein by reference in its entirety.

In another approach, the genome-derived single exon probes and microarrays of the present invention can be
15 used to interrogate genomic DNA, rather than pools of expressed message; this latter approach permits predisposition to and/or prognosis of diseases of bone marrow to be assessed through the massively parallel determination of altered copy number, deletion, or mutation
20 in the patient's genome of exons known to be expressed in human bone marrow. The algorithms set forth in WO 99/58720 can be applied to such genomic profiles without regard to the function of the protein encoded by the interrogated gene.

25 The utility is specific to the probe; at sufficiently high hybridization stringency, which stringencies are well known in the art - see Ausubel et al. and Maniatis et al. - each probe reports the level of expression of message specifically containing that ORF.

30 It should be appreciated, however, that the probes of the present invention, for which expression in the bone marrow has been demonstrated are useful for both measurement in the bone marrow and for survey of expression in other tissues.

35 Significant among such advantages is the presence

of probes for novel genes.

As mentioned above and further detailed in Examples 1 and 2, the methods described enable ORFs which are not present in existing expression databases to be
5 identified. And the fewer the number of tissues in which the ORF can be shown to be expressed, the more likely the ORF will prove to be part of a novel gene: as further discussed in Example 2, ORFs whose expression was measurable in only a single of the tested tissues were
10 represented in existing expression databases at a rate of only 11%, whereas 36% of ORFs whose expression was measurable in 9 tissues were present in existing expression databases, and fully 45% of those ORFs expressed in all ten tested tissues were present in existing expressed sequence
15 databases.

Either as tools for measuring gene expression or tools for surveying gene expression, the genome-derived single exon probes of the present invention have significant advantages over the cDNA or EST-based probes
20 that are currently available for achieving these utilities.

The genome-derived single exon probes of the present invention are useful in constructing genome-derived single exon microarrays; the genome-derived single exon microarrays, in turn, are useful devices for measuring and
25 for surveying gene expression in the human.

Gene expression analysis using microarrays – conventionally using microarrays having probes derived from expressed message – is well-established as useful in the biological research arts (see Lockhart et al. Nature 405,
30 827-836).

Microarrays have been used to determine gene expression profiles in cells in response to drug treatment (see, for example, Kaminski et al., "Global Analysis of Gene Expression in Pulmonary Fibrosis Reveals Distinct
35 Programs Regulating Lung Inflammation and Fibrosis," Proc.

Natl. Acad. Sci. USA 97(4):1778-83 (2000); Bartosiewicz et al., "Development of a Toxicological Gene Array and Quantitative Assessment of This Technology," *Arch. Biochem. Biophys.* 376(1):66-73 (2000)), viral infection (see for
5 example, Geiss et al., "Large-scale Monitoring of Host Cell Gene Expression During HIV-1 Infection Using cDNA Microarrays," *Virology* 266(1):8-16 (2000)) and during cell processes such as differentiation, senescence and apoptosis (see, for example, Shelton et al., "Microarray Analysis of
10 Replicative Senescence," *Curr. Biol.* 9(17):939-45 (1999); Voehringer et al., "Gene Microarray Identification of Redox and Mitochondrial Elements That Control Resistance or Sensitivity to Apoptosis," *Proc. Natl. Acad. Sci. USA* 97(6):2680-5 (2000)).

15 Microarrays have also been used to determine abnormal gene expression in diseased tissues (see, for example, Alon et al., "Broad Patterns of Gene Expression Revealed by Clustering Analysis of Tumor and Normal Colon Tissues Probed by Oligonucleotide Arrays," *Proc. Natl.*
20 *Acad. Sci. USA* 96(12):6745-50 (1999); Perou et al., "Distinctive Gene Expression Patterns in Human Mammary Epithelial Cells and Breast Cancers," *Proc. Natl. Acad. Sci. USA* 96(16):9212-7 (1999); Wang et al., "Identification of Genes Differentially Over-expressed in Lung Squamous Cell
25 Carcinoma Using Combination of cDNA Subtraction and Microarray Analysis," *Oncogene* 19(12):1519-28 (2000); Whitney et al., "Analysis of Gene Expression in Multiple Sclerosis Lesions Using cDNA Microarrays," *Ann. Neurol.* 46(3):425-8 (1999)), in drug discovery screens (see, for
30 example, Scherf et al., "A Gene Expression Database for the Molecular Pharmacology of Cancer," *Nat. Genet.* 24(3):236-44 (2000)) and in diagnosis to determine appropriate treatment strategies (see, for example, Sgroi et al., "In vivo Gene Expression Profile Analysis of Human Breast Cancer

Progression," *Cancer Res.* 59(22):5656-61 (1999)).

In microarray-based gene expression screens of pharmacological drug candidates upon cells, each probe provides specific useful data. In particular, it should be appreciated that even those probes that show no change in expression are as informative as those that do change, serving, in essence, as negative controls.

For example, where gene expression analysis is used to assess toxicity of chemical agents on cells, the failure of the agent to change a gene's expression level is evidence that the drug likely does not affect the pathway of which the gene's expressed protein is a part. Analogously, where gene expression analysis is used to assess side effects of pharmacological agents - whether in lead compound discovery or in subsequent screening of lead compound derivatives - the inability of the agent to alter a gene's expression level is evidence that the drug does not affect the pathway of which the gene's expressed protein is a part.

WO 99/58720 provides methods for quantifying the relatedness of a first and second gene expression profile and for ordering the relatedness of a plurality of gene expression profiles. The methods so described permit useful information to be extracted from a greater percentage of the individual gene expression measurements from a microarray than methods previously used in the art.

Other uses of microarrays are described in Gerhold et al., *Trends Biochem. Sci.* 24(5):168-173 (1999) and Zweiger, *Trends Biotechnol.* 17(11):429-436 (1999); Schena et al.

The invention particularly provides genome-derived single-exon probes known to be expressed in bone marrow. The individual single exon probes can be provided in the form of substantially isolated and purified nucleic acid, typically, but not necessarily, in a quantity

sufficient to perform a hybridization reaction.

Such nucleic acid can be in any form directly hybridizable to the message that contains the probe's ORF, such as double stranded DNA, single-stranded DNA
5 complementary to the message, single-stranded RNA complementary to the message, or chimeric DNA/RNA molecules so hybridizable. The nucleic acid can alternatively or additionally include either nonnative nucleotides, alternative internucleotide linkages, or both, so long as
10 complementary binding can be obtained. For example, probes can include phosphorothioates, methylphosphonates, morpholino analogs, and peptide nucleic acids (PNA), as are described, for example, in U.S. Patent Nos. 5,142,047; 5,235,033; 5,166,315; 5,217,866; 5,184,444; 5,861,250.

15 Usefully, however, such probes are provided in a form and quantity suitable for amplification, where the amplified product is thereafter to be used in the hybridization reactions that probe gene expression. Typically, such probes are provided in a form and quantity
20 suitable for amplification by PCR or by other well known amplification technique. One such technique additional to PCR is rolling circle amplification, as is described, *inter alia*, in U.S. Patent Nos. 5,854,033 and 5,714,320 and international patent publications WO 97/19193 and
25 WO 00/15779. As is well understood, where the probes are to be provided in a form suitable for amplification, the range of nucleic acid analogues and/or internucleotide linkages will be constrained by the requirements and nature of the amplification enzyme.

30 Where the probe is to be provided in form suitable for amplification, the quantity need not be sufficient for direct hybridization for gene expression analysis, and need be sufficient only to function as an amplification template, typically at least about 1, 10 or
35 100 pg or more.

Each discrete amplifiable probe can also be packaged with amplification primers, either in a single composition that comprises probe template and primers, or in a kit that comprises such primers separately packaged therefrom. As earlier mentioned, the ORF-specific 5' primers used for genomic amplification can have a first common sequence added thereto, and the ORF-specific 3' primers used for genomic amplification can have a second, different, common sequence added thereto, thus permitting, in this embodiment, the use of a single set of 5' and 3' primers to amplify any one of the probes. The probe composition and/or kit can also include buffers, enzyme, etc., required to effect amplification.

As mentioned earlier, when intended for use on a genome-derived single exon microarray of the present invention, the genome-derived single exon probes of the present invention will typically average at least about 100, 200, 300, 400 or 500 bp in length, including (and typically, but not necessarily centered about) the ORF. Furthermore, when intended for use on a genome-derived single exon microarray of the present invention, the genome-derived single exon probes of the present invention will typically not contain a detectable label.

When intended for use in solution phase hybridization, however — that is, for use in a hybridization reaction in which the probe is not first bound to a support substrate (although the target may indeed be so bound) — length constraints that are imposed in microarray-based hybridization approaches will be relaxed, and such probes will typically be labeled.

In such case, the only functional constraint that dictates the minimum size of such probe is that each such probe must be capable of specifically identifying in a hybridization reaction the exon from which it is drawn. In theory, a probe of as little as 17 nucleotides is capable

of uniquely identifying its cognate sequence in the human genome. For hybridization to expressed message – a subset of target sequence that is much reduced in complexity as compared to genomic sequence – even fewer nucleotides are
5 required for specificity.

Therefore, the probes of the present invention can include as few as 20, 25 or 50 bp or ORF, or more. In particular embodiments, the ORF sequences are given in SEQ ID NOS. 13,115 – 26,012, respectively, for probe SEQ ID
10 NOS. 1 – 13,114. The minimum amount of ORF required to be included in the probe of the present invention in order to provide specific signal in either solution phase or microarray-based hybridizations can readily be determined for each of ORF SEQ ID NOS. 13,115 – 26,012 individually
15 by routine experimentation using standard high stringency conditions.

Such high stringency conditions are described, *inter alia*, in Ausubel et al. and Maniatis et al. For microarray-based hybridization, standard high stringency
20 conditions can usefully be 50% formamide, 5X SSC, 0.2 µg/µl poly(dA), 0.2 µg/µl human c_{ot}1 DNA, and 0.5 % SDS, in a humid oven at 42°C overnight, followed by successive washes of the microarray in 1X SSC, 0.2% SDS at 55°C for 5 minutes, and then 0.1X SSC, 0.2% SDS, at 55°C for 20
25 minutes. For solution phase hybridization, standard high stringency conditions can usefully be aqueous hybridization at 65°C in 6X SSC. Lower stringency conditions, suitable for cross-hybridization to mRNA encoding structurally- and functionally-related proteins, can usefully be the same as
30 the high stringency conditions but with reduction in temperature for hybridization and washing to room temperature (approximately 25°C).

When intended for use in solution phase hybridization, the maximum size of the single exon probes
35 of the present invention is dictated by the proximity of

other expressed exons in genomic DNA: although each single exon probe can include intergenic and/or intronic material contiguous to the ORF in the human genome, each probe of the present invention will include portions of only one
5 expressed exon.

Thus, each single exon probe will include no more than about 25 kb of contiguous genomic sequence, more typically no more than about 20 kb of contiguous genomic sequence, more usually no more than about 15 kb, even more
10 usually no more than about 10 kb. Usually, probes that are maximally about 5 kb will be used, more typically no more than about 3 kb.

It will be appreciated that the Sequence Listing appended hereto presents, by convention, only that strand
15 of the probe and ORF sequence that can be directly translated reading from 5' to 3' end. As would be well understood by one of skill in the art, single stranded probes must be complementary in sequence to the ORF as present in an mRNA; it is well within the skill in the art
20 to determine such complementary sequence. It will further be understood that double stranded probes can be used in both solution-phase hybridization and microarray-based hybridization if suitably denatured.

Thus, it is an aspect of the present invention to
25 provide single-stranded nucleic acid probes that have sequence complementary to those described herein above and below, and double-stranded probes one strand of which has sequence complementary to the probes described herein.

The probes can, but need not, contain intergenic
30 and/or intronic material that flanks the ORF, on one or both sides, in the same linear relationship to the ORF that the intergenic and/or intronic material bears to the ORF in genomic DNA. The probes do not, however, contain nucleic acid derived from more than one expressed ORF.

35 And when intended for use in solution

hybridization, the probes of the present invention can usefully have detectable labels. Nucleic acid labels are well known in the art, and include, *inter alia*, radioactive labels, such as ^3H , ^{32}P , ^{33}P , ^{35}S , ^{125}I , ^{131}I ; fluorescent
5 labels, such as Cy3, Cy5, Cy5.5, Cy7, SYBR

Green and other labels described in Haugland, *Handbook of Fluorescent Probes and Research Chemicals*, 7th ed., Molecular Probes Inc., Eugene, OR (2000), or fluorescence resonance energy transfer tandem conjugates
10 thereof; labels suitable for chemiluminescent and/or enhanced chemiluminescent detection; labels suitable for ESR and NMR detection; and labels that include one member of a specific binding pair, such as biotin, digoxigenin, or the like.

15 The probes, either in quantity sufficient for hybridization or sufficient for amplification, can be provided in individual vials or containers.

Alternatively, such probes can usefully be packaged as a plurality of such individual genome-derived
20 single exon probes.

When provided as a collection of plural individual probes, the probes are typically made available in amplifiable form in a spatially-addressable ordered set, typically one per well of a microtiter dish. Although a 96
25 well microtiter plate can be used, greater efficiency is obtained using higher density arrays.

If, as earlier mentioned, the ORF-specific 5' primers used for genomic amplification had a first common sequence added thereto, and the ORF-specific 3'
30 primers used for genomic amplification had a second, different, common sequence added thereto, a single set of 5' and 3' primers can be used to amplify all of the probes from the amplifiable ordered set.

Such collections of genome-derived single exon
35 probes can usefully include a plurality of probes chosen

for the common attribute of expression in the human bone marrow.

In such defined subsets, typically at least 50, 60, 75, 80, 85, 90 or 95% or more of the probes will be
5 chosen by their expression in the defined tissue or cell type.

The single exon probes of the present invention, as well as fragments of the single exon probes comprising selectively hybridizable portions of the probe ORF, can be
10 used to obtain the full length cDNA that includes the ORF by (i) screening of cDNA libraries; (ii) rapid amplification of cDNA ends ("RACE"); or (iii) other conventional means, as are described, *inter alia*, in Ausubel et al. and Maniatis et al.

15 It is another aspect of the present invention to provide genome-derived single exon nucleic acid microarrays useful for gene expression analysis, where the term "microarray" has the meaning given in the definitional section of this description, *supra*.

20 The invention particularly provides genome-derived single-exon nucleic acid microarrays comprising a plurality of probes known to be expressed in human bone marrow. In preferred embodiments, the present invention provides human genome-derived single exon microarrays
25 comprising a plurality of probes drawn from the group consisting of SEQ ID NOS.: 1 - 13,114.

When used for gene expression analysis, the genome-derived single exon microarrays provide greater physical informational density than do the genome-derived
30 single exon microarrays that have lower percentages of probes known to be expressed commonly in the tested tissue. At a fixed probe density, for example, a given microarray surface area of the defined subset genome-derived single exon microarray can yield a greater number of expression
35 measurements. Alternatively, at a given probe density, the

same number of expression measurements can be obtained from a smaller substrate surface area. Alternatively, at a fixed probe density and fixed surface area, probes can be provided redundantly, providing greater reliability in
5 signal measurement for any given probe. Furthermore, with a higher percentage of probes known to be expressed in the assayed tissue, the dynamic range of the detection means can be adjusted to reveal finer levels discrimination among the levels of expression.

10 Although particularly described with respect to their utility as probes of gene expression, particularly as probes to be included on a genome-derived single exon microarray, each of the nucleic acids having SEQ ID NOS.: 1 - 13,114 contains an open-reading frame, set forth
15 respectively in SEQ ID NOS.: 13,115 - 26,012, that encodes a protein domain. Thus, each of SEQ ID NOS. 1 - 13,114 can be used, or that portion thereof in SEQ ID NOS. 13,115 - 26,012 used, to express a protein domain by standard in vitro recombinant techniques. See Ausubel et al. and
20 Maniatis et al.

Additionally, kits are available commercially that readily permit such nucleic acids to be expressed as protein in bacterial cells, insect cells, or mammalian cells, as desired (e.g., HAT™ Protein Expression &
25 Purification System, ClonTech Laboratories, Palo Alto, CA; Adeno-X™ Expression System, ClonTech Laboratories, Palo Alto, CA; Protein Fusion & Purification (pMAL™) System, New England Biolabs, Beverley, MA)

Furthermore, shorter peptides can be chemically
30 synthesized using commercial peptide synthesizing equipment and well known techniques. Procedures are described, *inter alia*, in Chan et al. (eds.), Fmoc Solid Phase Peptide Synthesis: A Practical Approach (Practical Approach Series, (Paper)), Oxford Univ. Press (March 2000) (ISBN:
35 0199637245); Jones, Amino Acid and Peptide Synthesis

(Oxford Chemistry Primers, No 7) , Oxford Univ. Press
(August 1992) (ISBN: 0198556683); and Bodanszky, Principles
of Peptide Synthesis (Springer Laboratory), Springer Verlag
(December 1993) (ISBN: 0387564314).

5 It is, therefore, another aspect of the invention
to provide peptides comprising an amino acid sequence
translated from SEQ ID NOS.: 13,115 - 26,012. Such amino
acid sequences are set out in SEQ ID NOS: 26,013 - 38,628.
Any such recombinantly-expressed or synthesized peptide of
10 at least 8, and preferably at least about 15, amino acids,
can be conjugated to a carrier protein and used to generate
antibody that recognizes the peptide. Thus, it is a
further aspect of the invention to provide peptides that
have at least 8, preferably at least 15, consecutive amino
15 acids.

The following examples are offered by way of
illustration and not by way of limitation.

20 EXAMPLE 1

Preparation of Single Exon Microarrays from ORFs Predicted
in Human Genomic Sequence

Bioinformatics Results

25 All human BAC sequences in fewer than 10 pieces
that had been accessioned in a five month period
immediately preceding this study were downloaded from
GenBank. This corresponds to ~2200 clones, totaling ~350
MB of sequence, or approximately 10% of the human genome.

30 After masking repetitive elements using the
program CROSS_MATCH, the sequence was analyzed for open
reading frames using three separate gene finding programs.
The three programs predict genes using independent
algorithmic methods developed on independent training sets:
35 GRAIL uses a neural network, GENEFINDER uses a hidden

Markoff model, and DICTION, a program proprietary to Genetics Institute, operates according to a different heuristic. The results of all three programs were used to create a prediction matrix across the segment of genomic DNA.

The three gene finding programs yielded a range of results. GRAIL identified the greatest percentage of genomic sequence as putative coding region, 2% of the data analyzed. GENEFINDER was second, calling 1%, and DICTION yielded the least putative coding region, with 0.8% of genomic sequence called as coding region.

The consensus data were as follows. GRAIL and GENEFINDER agreed on 0.7% of genomic sequence, GRAIL and DICTION agreed on 0.5% of genomic sequence, and the three programs together agreed on 0.25% of the data analyzed. That is, 0.25% of the genomic sequence was identified by all three of the programs as containing putative coding region.

ORFs predicted by any two of the three programs ("consensus ORFs") were assorted into "gene bins" using two criteria: (1) any 7 consecutive exons within a 25 kb window were placed together in a bin as likely contributing to a single gene, and (2) all ORFs within a 25 kb window were placed together in a bin as likely contributing to a single gene if fewer than 7 exons were found within the 25 kb window.

PCR

The largest ORF from each gene bin that did not span repetitive sequence was then chosen for amplification, as were all consensus ORFs longer than 500 bp. This method approximated one exon per gene; however, a number of genes were found to be represented by multiple elements.

Previously, we had determined that DNA fragments fewer than 250 bp in length do not bind well to the amino-

modified glass surface of the slides used as support substrate for construction of microarrays; therefore, amplicons were designed in the present experiments to approximate 500 bp in length.

5 Accordingly, after selecting the largest ORF per gene bin, a 500 bp fragment of sequence centered on the ORF was passed to the primer picking software, PRIMER3 (available online for use at <http://www-genome.wi.mit.edu/cgi-bin/primer/>). A first
10 additional sequence was commonly added to each ORF-unique 5' primer, and a second, different, additional sequence was commonly added to each ORF-unique 3' primer, to permit subsequent reamplification of the amplicon using a single set of "universal" 5' and 3' primers, thus immortalizing
15 the amplicon. The addition of universal priming sequences also facilitates sequence verification, and can be used to add a cloning site should some ORFs be found to warrant further study.

 The ORFs were then PCR amplified from genomic
20 DNA, verified on agarose gels, and sequenced using the universal primers to validate the identity of the amplicon to be spotted in the microarray.

 Primers were supplied by Operon Technologies (Alameda, CA). PCR amplification was performed by standard
25 techniques using human genomic DNA (Clontech, Palo Alto, CA) as template. Each PCR product was verified by SYBR® green (Molecular Probes, Inc., Eugene, OR) staining of agarose gels, with subsequent imaging by Fluorimager (Molecular Dynamics, Inc., Sunnyvale, CA). PCR
30 amplification was classified as successful if a single band appeared.

 The success rate for amplifying ORFs of interest directly from genomic DNA using PCR was approximately 75%.
FIG. 5 graphs the distribution of predicted ORF (exon)
35 length and distribution of amplified PCR products, with ORF

length shown in red and PCR product length shown in blue (which may appear black in the figure). Although the range of ORF sizes is readily seen to extend to beyond 900 bp, the mean predicted exon size was only 229 bp, with a median
5 size of 150 bp (n=9498). With an average amplicon size of 475 ± 25 bp, approximately 50% of the average PCR amplification product contained predicted coding region, with the remaining 50% of the amplicon containing either intron, intergenic sequence, or both.

10 Using a strategy predicated on amplifying about 500 bp, it was found that long exons had a higher PCR failure rate. To address this, the bioinformatics process was adjusted to amplify 1000, 1500 or 2000 bp fragments from exons larger than 500 bp. This improved the rate of
15 successful amplification of exons exceeding 500 bp, constituting about 9.2% of the exons predicted by the gene finding algorithms.

 Approximately 75% of the probes disposed on the array (90% of those that successfully PCR amplified) were
20 sequence-verified by sequencing in both the forward and reverse direction using MegaBACE sequencer (Molecular Dynamics, Inc., Sunnyvale, CA), universal primers, and standard protocols.

 Some genomic clones (BACs) yielded very poor PCR
25 and sequencing results. The reasons for this are unclear, but may be related to the quality of early draft sequence or the inclusion of vector and host contamination in some submitted sequence data.

 Although the intronic and intergenic material
30 flanking coding regions could theoretically interfere with hybridization during microarray experiments, subsequent empirical results demonstrated that differential expression ratios were not significantly affected by the presence of noncoding sequence. The variation in exon size was
35 similarly found not to affect differential expression

ratios significantly; however, variation in exon size was observed to affect the absolute signal intensity (data not shown).

The 350 MB of genomic DNA was, by the above-
5 described process, reduced to 9750 discrete probes, which were spotted in duplicate onto glass slides using commercially available instrumentation (MicroArray GenII Spotter and/or MicroArray GenIII Spotter, Molecular Dynamics, Inc., Sunnyvale, CA). Each slide additionally
10 included either 16 or 32 *E. coli* genes, the average hybridization signal of which was used as a measure of background biological noise.

Each of the probe sequences was BLASTed against the human EST data set, the NR data set, and SwissProt
15 GenBank (May 7, 1999 release 2.0.9).

One third of the probe sequences (as amplified) produced an exact match (BLAST Expect ("E") values less than $1 e^{-100}$) to either an EST (20% of sequences) or a known mRNA (13% of sequences). A further 22% of the probe
20 sequences showed some homology to a known EST or mRNA (BLAST E values from $1 e^{-5}$ to $1 e^{-99}$). The remaining 45% of the probe sequences showed no significant sequence homology to any expressed, or potentially expressed, sequences present in public databases.

25 All of the probe sequences (as amplified) were then analyzed for protein similarities with the SwissProt database using BLASTX, Gish et al., *Nature Genet.* 3:266 (1993). The predicted functional breakdowns of the 2/3 of probes identical or homologous to known sequences are
30 presented in Table 1.

Table 1

Function of Predicted ORFs As Deduced From Comparative Sequence Analysis
--

Total	V6 chip	V7 chip	Function Predicted from Comparative Sequence Analysis
211	96	115	Receptor
120	43	77	Zinc Finger
30	11	19	Homeobox
25	9	16	Transcription Factor
17	11	7	Transcription
118	57	61	Structural
95	39	56	Kinase
36	18	18	Phosphatase
83	31	52	Ribosomal
45	19	26	Transport
21	17	14	Growth Factor
17	12	5	Cytochrome
50	33	17	Channel

As can be seen, the two most common types of genes were transcription factors and receptors, making up 2.2% and 1.8% of the arrayed elements, respectively.

5

EXAMPLE 2

Gene Expression Measurements From Genome-Derived Single Exon Microarrays

10

The two genome-derived single exon microarrays prepared according to Example 1 were hybridized in a series of simultaneous two-color fluorescence experiments to (1)
 15 Cy3-labeled cDNA synthesized from message drawn individually from each of brain, heart, liver, fetal liver, placenta, lung, bone marrow, HeLa, BT 474, or HBL 100 cells, and (2) Cy5-labeled cDNA prepared from message

pooled from all ten tissues and cell types, as a control in each of the measurements. Hybridization and scanning were carried out using standard protocols and Molecular Dynamics equipment.

5 Briefly, mRNA samples were bought from commercial sources (Clontech, Palo Alto, CA and Amersham Pharmacia Biotech (APB)). Cy3-dCTP and Cy5-dCTP (both from APB) were incorporated during separate reverse transcriptions of 1 µg of polyA⁺ mRNA performed using 1 µg oligo(dT)12-18 primer
10 and 2 µg random 9mer primers as follows. After heating to 70°C, the RNA:primer mixture was snap cooled on ice. After snap cooling on ice, added to the RNA to the stated final concentration was: 1X Superscript II buffer, 0.01 M DTT, 100µM dATP, 100 µM dGTP, 100 µM dTTP, 50 µM dCTP, 50 µM
15 Cy3-dCTP or Cy5-dCTP 50 µM, and 200 U Superscript II enzyme. The reaction was incubated for 2 hours at 42°C. After 2 hours, the first strand cDNA was isolated by adding 1 U Ribonuclease H, and incubating for 30 minutes at 37°C. The reaction was then purified using a Qiagen PCR cleanup
20 column, increasing the number of ethanol washes to 5. Probe was eluted using 10 mM Tris pH 8.5.

Using a spectrophotometer, probes were measured for dye incorporation. Volumes of both Cy3 and Cy5 cDNA corresponding to 50 pmoles of each dye were then dried in a
25 Speedvac, resuspended in 30 µl hybridization solution containing 50% formamide, 5X SSC, 0.2 µg/µl poly(dA), 0.2 µg/µl human c_ot1 DNA, and 0.5 % SDS.

Hybridizations were carried out under a coverslip, with the array placed in a humid oven at 42°C
30 overnight. Before scanning, slides were washed in 1X SSC, 0.2% SDS at 55°C for 5 minutes, followed by 0.1X SSC, 0.2% SDS, at 55°C for 20 minutes. Slides were briefly dipped in water and dried thoroughly under a gentle stream of nitrogen.

35 Slides were scanned using a Molecular Dynamics

Gen3 scanner, as described. Schena (ed.), Microarray Biochip: Tools and Technology, Eaton Publishing Company/BioTechniques Books Division (2000) (ISBN: 1881299376).

5 Although the use of pooled cDNA as a reference permitted the survey of a large number of tissues, it attenuates the measurement of relative gene expression, since every highly expressed gene in the tissue/cell type-specific fluorescence channel will be present to a level of
10 at least 10% in the control channel. Because of this fact, both signal and expression ratios (the latter hereinafter, "expression" or "relative expression") for each probe were normalized using the average ratio or average signal, respectively, as measured across the whole slide.

15 Data were accepted for further analysis only when signal was at least three times greater than biological noise, the latter defined by the average signal produced by the *E. coli* control genes.

 The relative expression signal for these probes
20 was then plotted as function of tissue or cell type, and is presented in FIG. 6.

 FIG. 6 shows the distribution of expression across a panel of ten tissues. The graph shows the number of sequence-verified products that were either not
25 expressed ("0"), expressed in one or more but not all tested tissues ("1" - "9"), and expressed in all tissues tested ("10").

 Of 9999 arrayed elements on the two microarrays (including positive and negative controls and "failed"
30 products), 2353 (51%) were expressed in at least one tissue or cell type. Of the gene elements showing significant signal - where expression was scored as "significant" if the normalized Cy3 signal was greater than 1, representing signal 5-fold over biological noise (0.2) - 39% (991) were
35 expressed in all 10 tissues. The next most common class

(15%) consisted of gene elements expressed in only a single tissue.

The genes expressed in a single tissue were further analyzed, and the results of the analyses are
5 compiled in FIG. 7.

FIG. 7A is a matrix presenting the expression of all verified sequences that showed expression greater than 3 in at least one tissue. Each clone is represented by a column in the matrix. Each of the 10 tissues assayed is
10 represented by a separate row in the matrix, and relative expression of a clone in that tissue is indicated at the respective node by intensity of green shading, with the intensity legend shown in panel B. The top row of the matrix ("EST Hit") contains "bioinformatic" rather than
15 "physical" expression data — that is, presents the results returned by query of EST, NR and SwissProt databases using the probe sequence. The legend for "bioinformatic expression" (i.e., degree of homology returned) is presented in panel C. Briefly, white is known, black is
20 novel, with gray depicting nonidentical with significant homology (white: E values < 1e-100; gray: E values from 1e-05 to 1e-99; black: E values > 1e-05).

As FIG. 7 readily shows, heart and brain were demonstrated to have the greatest numbers of genes that
25 were shown to be uniquely expressed in the respective tissue. In brain, 200 uniquely expressed genes were identified; in heart, 150. The remaining tissues gave the following figures for uniquely expressed genes: liver, 100; lung, 70; fetal liver, 150; bone marrow, 75; placenta, 100;
30 HeLa, 50; HBL, 100; and BT474, 50.

It was further observed that there were many more "novel" genes among those that were up-regulated in only one tissue, as compared with those that were down-regulated in only one tissue. In fact, it was found that ORFs whose
35 expression was measurable in only a single of the tested

tissues were represented in sequencing databases at a rate of only 11%, whereas 36% of the ORFs whose expression was measurable in 9 of the tissues were present in public databases. As for those ORFs expressed in all ten tissues, 5 fully 45% were present in existing expressed sequence databases. These results are not unexpected, since genes expressed in a greater number of tissues have a higher likelihood of being, and thus of having been, discovered by EST approaches.

10

Comparison of Signal from Known and Unknown Genes

The normalized signal of the genes found to have high homology to genes present in the GenBank human EST database were compared to the normalized signal of those 15 genes not found in the GenBank human EST database. The data are shown in FIG. 8.

FIG. 8 shows the normalized Cy3 signal intensity for all sequence-verified products with a BLAST Expect ("E") value of greater than $1e-30$ (designated "unknown") 20 upon query of existing EST, NR and SwissProt databases, and shows in blue the normalized Cy3 signal intensity for all sequence-verified products with a BLAST Expect value of less than $1e-30$ ("known"). Note that biological background noise has an averaged normalized Cy3 signal intensity of 25 0.2.

As expected, the most highly expressed of the ORFs were "known" genes. This is not surprising, since very high signal intensity correlates with very commonly-expressed genes, which have a higher likelihood of being 30 found by EST sequence.

However, a significant point is that a large number of even the high expressers were "unknown". Since the genomic approach used to identify genes and to confirm their expression does not bias exons toward either the 3' 35 or 5' end of a gene, many of these high expression genes

will not have been detected in an end-sequenced cDNA library.

The significant point is that presence of the gene in an EST database is not a prerequisite for
5 incorporation into a genome-derived microarray, and further, that arraying such "unknown" exons can help to assign function to as-yet undiscovered genes.

Verification of Gene Expression

10 To ascertain the validity of the approach described above to identify genes from raw genomic sequence, expression of two of the probes was assayed using reverse transcriptase polymerase chain reaction (RT PCR) and northern blot analysis.

15 Two microarray probes were selected on the basis of exon size, prior sequencing success, and tissue-specific gene expression patterns as measured by the microarray experiments. The primers originally used to amplify the two respective ORFs from genomic DNA were used in RT PCR
20 against a panel of tissue-specific cDNAs (Rapid-Scan gene expression panel 24 human cDNAs) (OriGene Technologies, Inc., Rockville, MD).

Sequence AL079300_1 was shown by microarray hybridization to be present in cardiac tissue, and sequence
25 AL031734_1 was shown by microarray experiment to be present in placental tissue (data not shown). RT-PCR on these two sequences confirmed the tissue-specific gene expression as measured by microarrays, as ascertained by the presence of a correctly sized PCR product from the respective tissue
30 type cDNAs.

Clearly, all microarray results cannot, and indeed should not, be confirmed by independent assay methods, or the high throughput, highly parallel advantages of microarray hybridization assays will be lost. However,
35 in addition to the two RT-PCR results presented above, the

observation that 1/3 of the arrayed genes exist in expression databases provides powerful confirmation of the power of our methodology – which combines bioinformatic prediction with expression confirmation using genome-derived single exon microarrays – to identify novel genes from raw genomic data.

To verify that the approach further provides correct characterization of the expression patterns of the identified genes, a detailed analysis was performed of the microarrayed sequences that showed high signal in brain.

For this latter analysis, sequences that showed high (normalized) signal in brain, but which showed very low (normalized) signal (less than 0.5, determined to be biological noise) in all other tissues, were further studied. There were 82 sequences that fit these criteria, approximately 2% of the arrayed elements. The 10 sequences showing the highest signal in brain in microarray hybridizations are detailed in Table 2, along with assigned function, if known or reasonably predicted.

20

Table 2

Function of the Most Highly Expressed Genes Expressed Only in Brain				
Microarray Sequence Name	Normalized Signal	Expression Ratio	Homology to EST present in GenBank	Gene Function as described by GenBank
AP000217-1	5.2	+7.7	High	S-100 protein, b-chain, Ca ²⁺ binding protein expressed in central nervous

				system
AP000047-1	2.3		High	Unknown Function
AC006548-9	1.7		High	Similar to mouse membrane glyco-protein M6, expressed in central nervous system
AC007245-5	1.5		High	Similar to amphiphysin, a synaptic vesicle- associated protein. Ref 21
L44140-4	1.2	+2.0	High	Endothelial actin-binding protein found in nonmuscle filamin
AC004689-9	1.2	+3.5	High	Protein Phosphatase PP2A, neuronal/ downregulates activated protein kinases
AL031657-1	1.2	+3.0	High	Unknown function/ Contains the anhyrin motif, a common protein sequence motif
AC009266-2	1.1	+3.7	Low	Low homology to

				the Synaptotagmin I protein in rat/present at low levels throughout rat brain
AP000086-1	1.0	+2.7	Low	Unknown, very poor homology to collagen
AC004689-3	1.0		High	Protein Phosphatase PP2A, neuronal/ downregulates activated protein kinases

Of the ten sequences studied by these latter confirmatory approaches, eight were previously known. Of these eight, six had previously been reported to be
5 important in the central nervous system or brain. The exon giving the highest signal (AP00217-1) was found to be the gene encoding an S100B Ca²⁺ binding protein, reported in the literature to be highly and uniquely expressed in the central nervous system. Heizmann, *Neurochem. Res.* 9:1097
10 (1997).

A number of the brain-specific probe sequences (including AC006548-9, AC009266-2) did not have homology to any known human cDNAs in GenBank but did show homology to rat and mouse cDNAs. Sequences AC004689-9 and AC004689-3
15 were both found to be phosphatases present in neurons (Millward *et al.*, *Trends Biochem. Sci.* 24(5):186-191 (1999)). Two microarray sequences, AP000047-1 and AP000086-1 have unknown function, with AP000086-1 being absent from GenBank. Functionality can now be narrowed

down to a role in the central nervous system for both of these genes, showing the power of designing microarrays in this fashion.

Next, the function of the chip sequences with the highest (normalized) signal intensity in brain, regardless of expression in other tissues, was assessed. In this latter analysis, we found expression of many more common genes, since the sequences were not limited to those expressed only in brain. For example, looking at the 20 highest signal intensity spots in brain, 4 were similar to tubulin (AC00807905; AF146191-2; AC007664-4; AF14191-2), 2 were similar to actin (AL035701-2; AL034402-1), and 6 were found to be homologous to glyceraldehyde-3-phosphate dehydrogenase (GAPDH) (AL035604-1; Z86090-1; AC006064-L, AC006064-K; AC035604-3; AC006064-L). These genes are often used as controls or housekeeping genes in microarray experiments of all types.

Other interesting genes highly expressed in brain were a ferritin heavy chain protein, which is reported in the literature to be found in brain and liver (Joshi et al., *J. Neurol. Sci.* 134(Suppl):52-56 (1995)), a result duplicated with the array. Other highly expressed chip sequences included a translation elongation factor 1 α (AC007564-4), a DEAD-box homolog (AL023804-4), and a Y-chromosome RNA-binding motif (Chai et al., *Genomics* 49(2):283-89 (1998)) (AC007320-3). A low homology analog (AP00123-1/2) to a gene, DSCR1, thought to be involved in trisomy 21 (Down's syndrome), showed high expression in both brain and heart, in agreement with the literature (Fuentes et al., *Mol. Genet.* 4(10):1935-44 (1995)).

As a further validation of the approach, we selected the BAC AC006064 to be included on the array. This BAC was known to contain the GAPDH gene, and thus could be used as a control for the ORF selection process. The gene finding and exon selection algorithms resulted in

choosing 25 exons from BAC AC006064 for spotting onto the array, of which four were drawn from the GAPDH gene. Table 3 shows the comparison of the average expression ratio for the 4 exons from BAC006064 compared with the average expression ratio for 5 different dilutions of a commercially available GAPDH cDNA (Clontech).

Table 3

Comparison of Expression Ratio, for each tissue, of GAPDH		
	AC006064 (n = 4)	Control (n = 5)
Bone Marrow	-1.81 ± 0.11	-1.85 ± 0.08
Brain	-1.41 ± 0.11	-1.17 ± 0.05
BT474	1.85 ± 0.09	1.66 ± 0.12
Fetal Liver	-1.62 ± 0.07	-1.41 ± 0.05
HBL100	1.32 ± 0.05	2.64 ± 0.12
Heart	1.16 ± 0.09	1.56 ± 0.10
HeLa	1.11 ± 0.06	1.30 ± 0.15
Liver	-1.62 ± 0.22	-2.07 ±
Lung	-4.95 ± 0.93	-3.75 ± 0.21
Placenta	-3.56 ± 0.25	-3.52 ± 0.43

10

Each tissue shows excellent agreement between the experimentally chosen exons and the control, again demonstrating the validity of the present exon mining approach. In addition, the data also show the variability of expression of GAPDH within tissues, calling into question its classification as a housekeeping gene and utility as a housekeeping control in microarray experiments.

20 EXAMPLE 3

Representation of Sequence and Expression Data as a
"Mondrian"

For each genomic clone processed for microarray
5 as above-described, a plethora of information was
accumulated, including full clone sequence, probe sequence
within the clone, results of each of the three gene finding
programs, EST information associated with the probe
sequences, and microarray signal and expression for
10 multiple tissues, challenging our ability to display the
information.

Accordingly, we devised a new tool for visual
display of the sequence with its attendant annotation
which, in deference to its visual similarity to the
15 paintings of Piet Mondrian, is hereinafter termed a
"Mondrian". FIGS. 3 and 4 present the key to the
information presented on a Mondrian.

FIG. 9 presents a Mondrian of BAC AC008172 (bases
25,000 to 130,000 shown), containing the carbamyl phosphate
20 synthetase gene (AF154830.1). Purple background within the
region shown as field 81 in FIG. 3 indicates all 37 known
exons for this gene.

As can be seen, GRAIL II successfully identified
27 of the known exons (73%), GENEFINDER successfully
25 identified 37 of the known exons (100%), while DICTION
identified 7 of the known exons (19%).

Seven of the predicted exons were selected for
physical assay, of which 5 successfully amplified by PCR
and were sequenced. These five exons were all found to be
30 from the same gene, the carbamyl phosphate synthetase gene
(AF154830.1).

The five exons were arrayed, and gene expression
measured across 10 tissues. As is readily seen in the
Mondrian, the five chip sequences on the array show
35 identical expression patterns, elegantly demonstrating the

reproducibility of the system..

FIG. 10 is a Mondrian of BAC AL049839. We selected 12 exons from this BAC, of which 10 successfully sequenced, which were found to form between 5 and 6 genes. Interestingly, 4 of the genes on this BAC are protease inhibitors. Again, these data elegantly show that exons selected from the same gene show the same expression patterns, depicted below the red line. From this figure, it is clear that our ability to find known genes is very good. A novel gene is also found from 86.6 kb to 88.6 kb, upon which all the exon finding programs agree. We are confident we have two exons from a single gene since they show the same expression patterns and the exons are proximal to each other. Backgrounds in the following colors indicate a known gene (top to bottom):
red = kallistatin protease inhibitor (P29622);
purple = plasma serine protease inhibitor (P05154);
turquoise = α 1 anti-chymotrypsin (P01011); mauve = 40S ribosomal protein (P08865). Note that chip sequence 8 and 12 did not sequence verify.

EXAMPLE 4

Genome-Derived Single Exon Probes Useful For Measuring
Human Gene Expression

The protocols set forth in Examples 1 and 2, *supra*, were applied to additional human genomic sequence as it became newly available in GenBank to identify unique exons in the human genome that could be shown to be expressed at significant levels in bone marrow tissue.

These unique exons are within longer probe sequences. Each probe was completely sequenced on both strands prior to its use on a genome-derived single exon microarray; sequencing confirms the exact chemical

structure of each probe. An added benefit of sequencing is that it placed us in possession of a set of single base-incremented fragments of the sequenced nucleic acid, starting from the sequencing primer 3' OH. (Since the
5 single exon probes were first obtained by PCR amplification from genomic DNA, we were of course additionally in possession of an even larger set of single base incremented fragments of each of the 13,114 single exon probes, each fragment corresponding to an extension product from one of
10 the two amplification primers.)

The structures of the 13,114 unique single exon probes are clearly presented in the Sequence Listing as SEQ ID Nos.: 1 - 13,114 . The 16 nt 5' primer sequence and 16 nt 3' primer sequence present on the amplicon are not
15 included in the sequence listing. The sequences of the exons present within each of these probes is presented in the Sequence Listing as SEQ ID Nos.: 13,115 - 26,012, respectively. It will be noted that some amplicons have more than one exon, some exons are contained in more than
20 one amplicon.

As detailed in Example 2, expression was demonstrated by disposing the amplicons as single exon probes on nucleic acid microarrays and then performing two-color fluorescent hybridization analysis; significant
25 expression is based on a statistical confidence that the signal is significantly greater than negative biological control spots. The negative biological control is formed from spotted DNA sequences from a different species. Here, 32 sequences from E.Coli were spotted in duplicate to give
30 a total of 64 spots.

For each hybridisation (each slide, each colour) the median value of the signal from all of the spots is determined. The normalised signal value is the arithmetic mean of the signal from duplicate spots divided by the
35 population median.

Control spots are eliminated if there is more than a five-fold difference between each one of the duplicate spots raw signals.

The median of the signal from the remaining
5 control spots is calculated and all subsequent calculations are done with normalised signals.

Control spots having a signal of greater than median + 2.4 (the value 2.4 is roughly 12 times the observed standard deviation of control spot populations)
10 are eliminated. Spots with such high signals are considered to be "outliers".

The mean and standard deviation of the modified control spot populations are calculated.

The mean + 3x the standard deviation (mean +
15 (3*SD)) is used as the signal threshold qualifier for that particular hybridisation. Thus, individual thresholds are determined for each channel and each hybridisation.

This means that, assuming that the data is distributed normally, there is a 99% confidence that any
20 signal exceeding the threshold is significant.

The probes and their expression data are presented in Table 4, set forth respectively in Example 5. Example 5 presents the subset of probes that is significantly expressed in the human bone marrow and thus
25 presents the subset of probes that was recognized to be useful for measuring expression of their cognate genes in human bone marrow tissue.

The sequence of each of the exon probes identified by SEQ ID NOS.: 13,115 - 26,012 was individually
30 used as a BLAST (or, for SWISSPROT, BLASTX) query to identify the most similar sequence in each of dbEST, SwissProt (BLASTX), and NR divisions of GenBank. Because the query sequences are themselves derived from genomic sequence in GenBank, only nongenic hits from NR were
35 scored.

The smallest in value of the BLAST (or BLASTX) expect ("E") scores for each query sequence across the three database divisions was used as a measure of the "expression novelty" of the probe's ORF. Table 4 is sorted
5 in descending order based on this measure, reported as "Most Similar (top) Hit BLAST E Value". Those sequences for which no "Hit E Value" is listed are those exons which were found to have no similar sequences.

As sorted, Table 4 thus lists its respective
10 probes (by "AMPLICON SEQ ID NO.:" and additionally by the SEQ ID NO:. of the exon contained within the probe:"EXON SEQ ID NO.:") from least similar to sequences known to be expressed (i.e., highest BLAST E value), at the beginning of the table, to most similar to sequences known to be
15 expressed (i.e., lowest BLAST E value), at the bottom of the table.

Table 4 further provides, for each listed probe, the accession number of the database sequence that yielded the "Most Similar (top) Hit BLAST E Value", along with the
20 name of the database in which the database sequence is found ("Top Hit Database Source").

Table 4 further provides SEQ ID NOS.
corresponding to the predicted amino acid sequences where they have been determined for the probe and exon nucleotide
25 sequences. These are set out as PEPTIDE SEQ ID NOS.:. The peptide sequences for a given exon are predicted as follows: Since each chip exon is a consensus sequence drawn from predictions from various exon finding programs (i.e. Grail, GeneFinder and GenScan), the multiple initial ORFs
30 are first determined in a uniform way according to each prediction. In particular, the reading frame for predicting the first amino acid in the peptide sequence always starts with the first base of any codon and ends with the last base of non-termination codon. Next, for each strand of the
35 exon, initial ORFs are merged into one or more final ORFs

in an exhaustive process based on the following criteria:

1) the merging ORFs must be overlapping, and 2) the merging ORFs must be in the same frame.

The Sequence Listing, which is a superset of all
5 of the data presented in Table 4, further includes, for each probe, the most similar hit, with accession number and BLAST E value, from the each of the three queried databases.

Table 4 further lists, for each probe, a portion
10 of the descriptor for the top hit ("Top Hit Descriptor") as provided in the sequence database. For those ORFs that are similar in sequence, but nonidentical to known sequences (e.g., those with BLAST E values between about $1e-05$ and $1e-100$), the descriptor reveals the likely function of the
15 protein encoded by the probe's ORF.

Using BLAST E value cutoffs of $1e-05$ (i.e., 1×10^{-5}) and $1e-100$ (i.e., 1×10^{-100}) as evidence of similarity to sequences known to be expressed is of course arbitrary: in Example 2, *supra*, a BLAST E value of $1e-30$ was used as
20 the boundary when only two classes were to be defined for analysis (unknown, $>1e-30$; known $<1e-30$) (see also FIG. 8). Furthermore, even when the "Most Similar (Top) Hit BLAST E Value" is low, e.g., less than about $1e-100$ — which is probative evidence that the query sequence has previously
25 been shown to be expressed — the top hit is highly unlikely exactly to match the probe sequence.

First, such expression entries typically will not have the intronic and/or intergenic sequence present within the single exon probes listed in the Table. Second, even
30 the ORF itself is unlikely in such cases to be present identically in the databases, since most of the EST and mRNA clones in existing databases include multiple exons, without any indication of the location of exon boundaries.

As noted, the data presented in Table 4 represent
35 a proper subset of the data present within the attached

sequence listing. For each amplicon probe (SEQ ID NOs.: 1 - 13,114) and probe exon (SEQ ID NOs.: 13,115 - 26,012, respectively), the sequence listing further provides, through iterated annotation fields <220> and <223>:

5 (a) the accession number of the BAC from which the sequence was derived ("MAP TO"), thus providing a link to the chromosomal map location and other information about the genomic milieu of the probe sequence;

(b) the most similar sequence provided by BLAST
10 query of the EST database, with accession number and BLAST E value for the "hit";

(c) the most similar sequence provided by BLAST query of the GenBank NR database, with accession number and BLAST E value for the "hit"; and

15 (d) the most similar sequence provided by BLASTX query of the SWISSPROT database, with accession number and BLAST E value for the "hit".

20 EXAMPLE 5

Genome-Derived Single Exon Probes Useful For Measuring Expression of Genes in Human Bone marrow

Table 4 (546 pages) presents expression, homology, and
25 functional information for the genome-derived single exon probes that are expressed significantly in human bone marrow.

30

CLAIMS

1. A spatially-addressable set of single exon nucleic acid probes for measuring gene expression in a sample derived
5 from human bone marrow comprising a plurality single exon nucleic probes, said probes comprising any one of the nucleotide sequences set out in SEQ ID NOS: 1 - 13,114 or a complementary sequence, or a portion of such a sequence.
- 10 2. A spatially-addressable set of single exon nucleic acid probes as claimed in claim 1 wherein each of said plurality of probes is separately and addressably amplifiable.
3. A spatially-addressable set of single exon nucleic acid
15 probes as claimed in claim 1 wherein each of said plurality of probes is separately and addressably isolatable from said plurality.
4. A spatially-addressable set of single exon nucleic acid
20 probes as claimed in any of claims 1 to 3 wherein said probes comprise any one of the nucleotide sequences set out in SEQ ID NOS.: 13,115 - 26,012.
5. A spatially-addressable set of single exon nucleic acid
25 probes as claimed in any of claims 1 to 4, wherein each of said plurality of probes is amplifiable using at least one common primer.
6. A spatially-addressable set of single exon nucleic acid
30 probes as claimed in any of claims 1 to 5 wherein the set comprises between 50 - 20,000 single exon nucleic acid probes.
7. A spatially-addressable set of single exon nucleic acid
35 probes as claimed in any of claims 1 to 6, wherein the

average length of the single exon nucleic acid probes is between 200 and 500 bp.

8. A spatially-addressable set of single exon nucleic acid probes as claimed in any of claims 1 to 7, wherein at least 50% of said single exon nucleic acid probes lack prokaryotic and bacteriophage vector sequence.

9. A spatially-addressable set of single exon nucleic acid probes as claimed in any of claims 1 to 8, wherein at least 50% of said single exon nucleic acid probes lack homopolymeric stretches of A or T.

10. A spatially-addressable set of single exon nucleic acid probes as claimed in any of claims 1 - 9 characterised in that said set of probes is addressably disposed upon a substrate.

11. A spatially-addressable set of single exon nucleic acid probes as claimed in claim 10 wherein said substrate is selected from glass, amorphous silicon, crystalline silicon and plastic.

12. A microarray comprising a spatially addressable set of single exon nucleic acid probes as claimed in any of claims 1 - 11.

13. A single exon nucleic acid probe for measuring human gene expression in a sample derived from human bone marrow comprising a nucleotide sequence as set out in any of SEQ ID NOs.: 1 - 13,114 or a complementary sequence or a fragment thereof wherein said probe hybridizes at high stringency to a nucleic acid molecule expressed in the human bone marrow.

14. A single exon nucleic acid probe as claimed in claim 13 comprising a nucleotide sequence as set out in any of SEQ ID NOs.: 13,115 - 26,012 or a complementary sequence or a fragment thereof.

5

15. A single exon nucleic acid probe for measuring human gene expression in a sample derived from human bone marrow which is a nucleic acid molecule having a sequence encoding a peptide comprising a peptide sequence as set out in any
10 of SEQ ID NOs.: 26,013 - 38,628, or a complementary sequence or a fragment thereof wherein said probe hybridizes at high stringency to a nucleic acid expressed in the human bone marrow.

15 16. A single exon nucleic acid probe as claimed in any one of claims 13 to 15 wherein said single exon nucleic acid probe comprises between 15 and 25 contiguous nucleotides of said SEQ ID NO.

20 17. A single exon nucleic acid probe as claimed in any one of claims 13 to 15, wherein said probe is between 3 - 25 kb in length.

18. A single exon nucleic acid probe as claimed in any one
25 of claims 13 - 17, wherein said probe is DNA, RNA or PNA.

19. A single exon nucleic acid probe as claimed in any one of claims 13 - 18, wherein said probe is detectably labeled.

30

20. A single exon nucleic acid probe as claimed in any one of claims 13 - 19, wherein said probe lacks prokaryotic and bacteriophage vector sequence.

35 21. A single exon nucleic acid probe as claimed in any one

of claims 13 - 20, wherein said probe lacks homopolymeric stretches of A or T.

22. A method of measuring gene expression in a sample
5 derived from human bone marrow, comprising:
 contacting the microarray of claim 12, with a first
 collection of detectably labeled nucleic acids,
 said first collection of nucleic acids derived
 from mRNA of human bone marrow; and then
10 measuring the label detectably bound to each probe of
 said microarray.

23. A method of identifying exons in a eukaryotic genome,
comprising:
15 algorithmically predicting at least one exon from
 genomic sequence of said eukaryote; and then
 detecting specific hybridization of detectably labeled
 nucleic acids to a single exon probe,
wherein said detectably labeled nucleic acids are derived
20 from mRNA from the bone marrow of said eukaryote, said
probe is a single exon probe having a fragment identical in
sequence to, or complementary in sequence to, said
predicted exon, said probe is included within a microarray
according to claim 12, and said fragment is selectively
25 hybridizable at high stringency.

24. A method of assigning exons to a single gene,
comprising:
 identifying a plurality of exons from genomic
30 sequence according to the method of claim 23; and
 then
 measuring the expression of each of said exons in a
 plurality of tissues and/or cell types using
 hybridization to single exon microarrays having a
35 probe with said exon,

wherein a common pattern of expression of said exons in said plurality of tissues and/or cell types indicates that the exons should be assigned to a single gene.

- 5 25. A nucleic acid sequence as set out in any of SEQ ID Nos: 1 - 26,012 which encodes a peptide.

26. A peptide encoded by a sequence as set out in any of SEQ ID Nos: 1 - 26,012.

10

27. A peptide comprising a sequence as set out in any of SEQ ID Nos: 26,013 - 38,628.

Page 1 of 546
Table 4
Single Exon Probes Expressed in Bone Marrow

Probe SEQ ID NO:	Exon SEQ ID NO:	ORF SEQ ID NO:	Expression Signal	Most Similar (Top) Hit BLAST E Value	Top Hit Accession No.	Top Hit Database Source	Top Hit Descriptor
450	13523	28455	4.57				
890	13945	26903	10.46				
1046	14062		3.88				
1305	14341	27305	7.19				
1618	14650	27628	4.21				
1842	14674	27847	6.05				
1738	14768	27753	2.88				
1764	14783	27778	1.27				
1770	14789	27785	8.04				
1608	14832	27928	1.24				
1885	15016	28023	2.44				
2175	15191	28212	2.92				
2287	15300	28324	2.97				
3200	16255	28175	3.13				
3484	16510	28431	1.32				
3527	16573	28498	10.05				
3574	16619		0.85				
3868	17008		1.15				
4225	17254	30141	1.66				
4290	17319	30188	6.25				
4310	17339	30218	0.83				
4310	17339	30219	0.83				
4384	17391		1.03				
4420	17447	30338	0.83				
4874	17891	30780	1.27				
4959	17874	30855	0.74				
5083	18093	30909	5.95				
5085	18105	30980	1.42				
5329	18435	31187	1.78				
5329	18435	31188	1.78				
5496	18596		4.07				
5678	18773		7.77				
5782	18586		3.48				
5824	18914	32097	0.65				

Page 2 of 546
Table 4
Single Exon Probes Expressed in Bone Marrow

Probe SEQ ID NO:	Exon SEQ ID NO:	ORF SEQ ID NO:	Expression Signal	Most Similar (Top) Hit BLAST E Value	Top Hit Accession No.	Top Hit Database Source	Top Hit Descriptor
5830	18920	32103	3.19				
6138	25655	32443	1.68				
6167	19242	32473	2.11				
6556	19616		1.24				
6700	19757	33034	0.89				
6700	19757	33035	0.89				
7332	20303	33847	1.52				
7332	20303	33848	1.52				
7642	20602	33966	1.45				
7642	20602	33967	1.45				
8114	21051		0.61				
8386	21365	34774	1.55				
8630	21797	35218	1.21				
9212	22178	35608	0.57				
9212	22178	35609	0.57				
9892	22845	36302	5.61				
10124	23050	36529	0.69				
10241	23166	36653	1.44				
10363	23305	36762	0.91				
10675	23597	37093	0.49				
10675	23597	37094	0.49				
10784	23715	37216	0.6				
10784	23715	37217	0.6				
11043	24007		2.14				
11368	24314		1.61				
11687	24653	38232	1.82				
11829	24712		1.94				
12600	25302		1.5				
12688	25476	31730	1.34				
6170	19246	32477	15.3	9.9E+00	AJ239028.1	NT	Homo sapiens LSS gene, partial, exons 15, 16, 17 and 18
8839	21308	34723	1.75	9.8E+00	U32716.1	NT	Haemophilus influenzae Rd section 31 of 163 of the complete genome
10100	23026	36502	0.47	9.8E+00	Y18930.1	NT	Sulfolobus solfataricus 281 kb genomic DNA fragment, strain P2
10100	23026	36503	0.47	9.8E+00	Y18930.1	NT	Sulfolobus solfataricus 281 kb genomic DNA fragment, strain P2
7194	20218	33549	0.66	9.6E+00	AF065630.1	NT	Gallus gallus ornithine transcarbamylase (OTC) gene, exon 1

Page 3 of 546
Table 4

Single Exon Probes Expressed in Bone Marrow

Probe SEQ ID NO.	Exon SEQ ID NO.	ORF SEQ ID NO.	Expression Signal	Most Similar (Top) Hit BLAST E Value	Top Hit Accession No.	Top Hit Database Source	Top Hit Descriptor
7194	20218	33550	0.83	9.6E+00	AF065630.1	NT	Gallus gallus ornithine transcarbamylase (OTC) gene, exon 1
10787	23708	37208	1.19	9.6E+00	AF242432.1	NT	Mus musculus Naip3 gene, exon 1; neuronal apoptosis inhibitory protein 1 (Naip1) and general transcription factor I1H polypeptide 2 (Gt2h2) genes, complete cds
10787	23708	37210	1.19	9.6E+00	AF242432.1	NT	Mus musculus Naip3 gene, exon 1; neuronal apoptosis inhibitory protein 1 (Naip1) and general transcription factor I1H polypeptide 2 (Gt2h2) genes, complete cds
2935	15993	28913	3.21	9.4E+00	AB043785.1	NT	Mus musculus AT3 gene for antithrombin, complete cds
8457	19522	32773	0.54	9.4E+00	P75130	SWISSPROT	HYPOTHETICAL PROTEIN MG447 HOMOLOG
11911	24792	38381	2.45	9.4E+00	O99825	SWISSPROT	NADHUBIQUINONE OXIDOREDUCTASE CHAIN 4
11911	24792	38382	2.45	9.4E+00	O99825	SWISSPROT	NADHUBIQUINONE OXIDOREDUCTASE CHAIN 4
8435	21404	34817	0.86	9.3E+00	AF130990.1	NT	Homo sapiens ectodysplasin-A receptor protein (EDAR) gene, exons 2, 3, and 4
8358	22321	35748	3.15	9.3E+00	P11210	SWISSPROT	IMMEDIATE-EARLY PROTEIN 1 (IE1) (IMMEDIATE-EARLY PHOSPHOPROTEIN PP89)
5368	18474	31346	2.63	9.1E+00	AF095609.1	NT	Leuciscus cephalus orientalis cytochrome b (cyt b) gene, partial cds; mitochondrial gene for mitochondrial product
5368	18474	31347	2.63	9.1E+00	AF095609.1	NT	Leuciscus cephalus orientalis cytochrome b (cyt b) gene, partial cds; mitochondrial gene for mitochondrial product
9785	22726		0.95	9.0E+00	P09241	SWISSPROT	RHODOPSIN
6152	19227	32456	5.62	8.9E+00	BE971806.1	EST_HUMAN	601651039R1 NIH_MGC_91 Homo sapiens cDNA clone IMAGE:3934562 3'
6517	19580	32837	2.21	8.7E+00	AB019788.1	NT	Cynops pyrrhogaster Cptb33 premature mRNA, partial cds
6517	19580	32838	2.21	8.7E+00	AB019788.1	NT	Cynops pyrrhogaster Cptb33 premature mRNA, partial cds
440	13514	26445	1.03	8.4E+00	5031804	NT	Homo sapiens insulin receptor substrate 1 (IRS1) mRNA
9810	21133	34536	3.58	8.1E+00	AJ131719.1	NT	Zea mays mRNA for legumain-like protease (ses2a)
11504	24446		1.54	8.0E+00	P41820	SWISSPROT	BREFELDIN A RESISTANCE PROTEIN
8491	21459		0.82	7.6E+00	Z21488.1	NT	African swine fever virus NP1450L gene encoding RNA polymerase largest subunit
7599	20632		2.17	7.5E+00	AL445065.1	NT	Thermoplasma acidophilum complete genome; segment 3/5
8704	21672	35095	1.58	7.5E+00	P35441	SWISSPROT	THROMBOSPONDIN 1 PRECURSOR
8704	21672	35098	1.58	7.5E+00	P35441	SWISSPROT	THROMBOSPONDIN 1 PRECURSOR
5898	18985	32178	3.53	7.4E+00	BF700517.1	EST_HUMAN	602128879F1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:4285506 5'
9108	22072	35498	2.72	7.4E+00	P04929	SWISSPROT	HISTIDINE-RICH GLYCOPROTEIN PRECURSOR
9108	22072	35499	2.72	7.4E+00	P04929	SWISSPROT	HISTIDINE-RICH GLYCOPROTEIN PRECURSOR
2990	16048	28968	4.7	7.2E+00	L12051.1	NT	Lycopodium obscurum Mill. GTPase (SAR2) mRNA, complete cds
2990	16048	28969	4.7	7.2E+00	L12051.1	NT	Lycopodium obscurum Mill. GTPase (SAR2) mRNA, complete cds
7230	20262	33586	1.12	7.2E+00	BE179090.1	EST_HUMAN	RC0-HT0613-200300-031-407 HT0613 Homo sapiens cDNA
7356	20326	33673	1.1	7.1E+00	P28168	SWISSPROT	ZINC-FINGER PROTEIN 1 (ZINC-FINGER HOMEODOMAIN PROTEIN 1)
7356	20326	33674	1.1	7.1E+00	P28168	SWISSPROT	ZINC-FINGER PROTEIN 1 (ZINC-FINGER HOMEODOMAIN PROTEIN 1)

Page 4 of 546
Table 4

Single Exon Probes Expressed in Bone Marrow

Probe SEQ ID NO:	Exon SEQ ID NO:	ORF SEQ ID NO:	Expression Signal	Most Similar (Top) Hit BLAST E Value	Top Hit Accession No.	Top Hit Database Source	Top Hit Descriptor
9957	22884		6.86	7.1E+00	AL161595.2	NT	Arabidopsis thaliana DNA chromosome 4, contig fragment No. 91
11715	24678	38256	2.96	7.1E+00	P06850	SWISSPROT	HYPOTHETICAL 17.3 KDA PROTEIN IN MRDA-PHPB INTERGENIC REGION
11809	24790	38379	4.81	7.1E+00	P06105	SWISSPROT	ME17 PROTEIN [INCLUDES: O-ACETYLHOMOSERINE SULFHYDRYLASE (OAH SULFHYDRYLASE);
10341	23265	36744	3.43	7.0E+00	P48610	SWISSPROT	O-ACETYL-SERINE SULFHYDRYLASE (OAS SULFHYDRYLASE)]
11580	24818	38074	1.7	7.0E+00	O22469	SWISSPROT	ARGININE KINASE (AK)
8628	21594	35014	3.94	6.9E+00	P35679	SWISSPROT	WD-40 REPEAT PROTEIN MS3
10718	23638	37131	1.32	6.9E+00	P44834	SWISSPROT	60S RIBOSOMAL PROTEIN L4 (L2)
10734	23658	37149	0.44	6.9E+00	P34226	SWISSPROT	DNA MISMATCH REPAIR PROTEIN MUTS
8240	21209	34613	1.31	6.8E+00	W03412.1	EST_HUMAN	SKT6 PROTEIN
8240	21209	34614	1.31	6.8E+00	W03412.1	EST_HUMAN	z807c11.11 Soares melanocyte 2Nblm Homo sapiens cDNA clone IMAGE:291860 5'
							z807c11.11 Soares melanocyte 2Nblm Homo sapiens cDNA clone IMAGE:291860 5'
9488	22452		1.35	6.8E+00	P36307	SWISSPROT	OUTER CAPSID PROTEIN VP4 (HEMAGGLUTININ) (OUTER LAYER PROTEIN VP4) [CONTAINS:
10569	23401	36883	3.31	6.8E+00	Q03570	SWISSPROT	OUTER CAPSID PROTEINS VP6 AND VP8]
5356	18461		0.74	6.8E+00	Q98028	SWISSPROT	HYPOTHETICAL 157.0 KDA PROTEIN C38G10.5 IN CHROMOSOME III
6695	19752	33029	0.72	6.8E+00	BF672121.1	EST_HUMAN	CATECHOL-O-METHYLTRANSFERASE, SOLUBLE FORM (S-COMT)
10434	23356	36842	1.87	6.8E+00	Q9ZE07	SWISSPROT	602162573F1 NIH_MGC_81 Homo sapiens cDNA clone IMAGE:4283427 5'
11463	24406		1.87	6.8E+00	Q9ZE07	SWISSPROT	URIDYLATE KINASE (UK) (URIDINE MONOPHOSPHATE KINASE) (UMP KINASE)
9535	22498	35946	2.49	6.6E+00	Q10309	SWISSPROT	URIDYLATE KINASE (UK) (URIDINE MONOPHOSPHATE KINASE) (UMP KINASE)
10667	23589	37089	7.17	6.5E+00	P03374	SWISSPROT	PROBABLE CATION-TRANSPORTING ATPASE C8C3.05C
10099	23025	36501	0.47	6.5E+00	BE66001.1	EST_HUMAN	ENV POLYPROTEIN [CONTAINS: COAT PROTEIN GP82; COAT PROTEIN GP36]
10927	23947	37362	1.17	6.2E+00	AY010901.1	EST_HUMAN	601678435F1 NIH_MGC_53 Homo sapiens cDNA clone IMAGE:3960869 5'
7236	20257	33591	0.55	6.2E+00	6754621	NT	Schizaphyllum commune unknown mRNA
10175	23100	36580	1.35	6.0E+00	BE780163.1	EST_HUMAN	Mus musculus mannosidase 2, alpha B1 (Man2b1), mRNA
10880	23800	37302	0.46	6.0E+00	AP000006.1	NT	601468031F1 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:3871303 5'
10880	23800	37303	0.63	6.0E+00	AE001862.1	NT	Pyrococcus horikoshii OT3 genomic DNA, 1166001-1485000 nt. position (8/7)
			0.63	6.0E+00	AE001862.1	NT	Deinococcus radiodurans R1 section 1 of 2 of the complete chromosome 2
						NT	Deinococcus radiodurans R1 section 1 of 2 of the complete chromosome 2
						NT	Mus musculus mixed lineage kinase 3 (Mlk3) and two pore domain K+ channel subunit (Kcnk8) genes,
6670	19727	33003	6.7	5.9E+00	AF155142.1	NT	complete cds
3538	16582		0.88	5.8E+00	7661557	NT	Homo sapiens DESG1 protein (DESG1), mRNA
7369	20339	33680	0.65	5.7E+00	AF302046.1	NT	Mus musculus immunoglobulin scavenger receptor IgSR mRNA, complete cds
7369	20339	33681	0.65	5.7E+00	AF302046.1	NT	Mus musculus immunoglobulin scavenger receptor IgSR mRNA, complete cds
7819	20768		1.5	5.6E+00	P75080	SWISSPROT	DNA POLYMERASE III, ALPHA CHAIN POLC-TYPE (POLII)
11805	23960	37485	2.98	5.6E+00	Q56276	SWISSPROT	LYCOPENE BETA CYCLASE

Table 4

Single Exon Probes Expressed in Bone Marrow

Probe SEQ ID NO:	Exon SEQ ID NO:	ORF SEQ ID NO:	Expression Signal	Most Similar (Top) Hit BLAST E Value	Top Hit Accession No.	Top Hit Database Source	Top Hit Descriptor
6380	10448	32680	0.78	5.5E+00	P47447	SWISSPROT	HEAT-INDUCIBLE TRANSCRIPTION REPRESSOR HRC
10138	23084		0.47	5.5E+00	P13983	SWISSPROT	EXTENSIN PRECURSOR (CELL WALL HYDROXYPROLINE-RICH GLYCOPROTEIN)
11803	23938	37482	2.35	5.5E+00	P11980	SWISSPROT	PNEUMOLYSIN (THIOL-ACTIVATED CYTOLYSIN)
7115	20049	33351	1.1	5.4E+00	X02212.1	NT	Chicken alpha-cardiac actin gene
7115	20049	33362	1.1	5.4E+00	X02212.1	NT	Chicken alpha-cardiac actin gene
7552	20515		0.84	5.4E+00	Q88435	SWISSPROT	NEL PROTEIN PRECURSOR (NEL-RELATED PROTEIN 2)
8143	21080	34480	0.6	5.4E+00	P90391	SWISSPROT	NEUROPEPTIDE Y RECEPTOR TYPE 4 (NPY4-R) (PANCREATIC POLYPEPTIDE RECEPTOR 1) (PP1)
							VITELLOGENIN PRECURSOR (VTG) [CONTAINS: LIPOVITELLIN LV-1N; LIPOVITELLIN LV-1C;
8202	21172		1.72	5.4E+00	Q91082	SWISSPROT	LIPOVITELLIN LV-2J
9151	22117	35543	0.73	5.4E+00	P40379	SWISSPROT	REPI PROTEIN
9151	22117	35544	0.73	5.4E+00	P40379	SWISSPROT	REPI PROTEIN
10396	23318	38788	1.41	5.4E+00	Q17094	SWISSPROT	RHODOPSIN
10396	23318	38800	1.41	5.4E+00	Q17094	SWISSPROT	RHODOPSIN
4824	17841	30739	1.22	5.3E+00	L43128.1	NT	Bovine immunodeficiency-like virus surface envelope gene, 5' end of cds
6635	19693		0.6	5.3E+00	P41779	SWISSPROT	HOMEOBOX PROTEIN CEH-20
8415	21384		3.9	5.3E+00	P54098	SWISSPROT	DNA POLYMERASE GAMMA (MITOCHONDRIAL DNA POLYMERASE CATALYTIC SUBUNIT)
9335	22300		0.53	5.3E+00	AB034990.1	NT	Homo sapiens HERPUD1 gene for stress protein Herp, complete cds
11950	24829	39425	1.84	5.3E+00	Q27905	SWISSPROT	PROBABLE ANTIBACTERIAL PEPTIDE POLYPROTEIN PRECURSOR
12093	24964	39559	2.34	5.3E+00	Z72863.1	NT	S.cerevisiae chromosome VII reading frame ORF YGL141W
12093	24964	39560	2.34	5.3E+00	Z72863.1	NT	S.cerevisiae chromosome VII reading frame ORF YGL141W
5539	18638		1.22	5.2E+00	BE184840.1	EST_HUMAN	QV4-HT0681-270400-186-09 HT0681 Homo sapiens cDNA
10738	23660		0.78	5.2E+00	AF248070.1	NT	Drosophila orientacea R1B retrotransposable element reverse transcriptase gene, partial cds
11527	24468		1.46	5.2E+00	Q10136	SWISSPROT	HYPOTHETICAL 61.1 KD PROTEIN C29E2.03C IN CHROMOSOME 1
9313	22278	35709	0.85	5.1E+00	O16005	SWISSPROT	RHODOPSIN
10184	23109	36592	1.07	5.1E+00	P09182	SWISSPROT	COLICIN N IMMUNITY PROTEIN (MICROGIN N IMMUNITY PROTEIN)
11617	24555	38117	3.01	5.1E+00	P55200	SWISSPROT	ZINC FINGER PROTEIN HRX (ALL-1)
6418	19485	32734	0.65	5.0E+00	BF310443.1	EST_HUMAN	601894910F1 NIH_MGC_19 Homo sapiens cDNA clone IMAGE:4124114 5'
10554	23476		0.65	5.0E+00	BF308561.1	EST_HUMAN	601890420F1 NIH_MGC_17 Homo sapiens cDNA clone IMAGE:4131509 5'
10796	23717	37219	3.78	5.0E+00	AF162445.2	NT	Canis familiaris skeletal muscle chloride channel CIC-1 (CLCN1) mRNA, complete cds
11622	24560	38122	6.39	5.0E+00	Z83860.1	NT	Mycobacterium tuberculosis H37Rv complete genome, segment 103/162
							Human hereditary haemochromatosis region, histone 2A-like protein gene, hereditary haemochromatosis (HLA-H) gene, RoRet gene, and sodium phosphate transporter (NPT3) gene, complete cds
10592	23514		0.63	4.9E+00	U91328.1	NT	Emura australis histone H3 (H3) gene, partial cds
4090	17124		9.81	4.8E+00	AF185255.1	NT	

Page 6 of 546
Table 4

Single Exon Probes Expressed in Bone Marrow

Probe SEQ ID NO:	Exon SEQ ID NO:	ORF SEQ ID NO:	Expression Signal	Most Similar (Top) Hit BLAST E Value	Top Hit Accession No.	Top Hit Database Source	Top Hit Descriptor
8494	21462	34878	0.53	4.8E+00	BF367908.1	EST_HUMAN	RC9-GN0042-100800-011-g10 GN0042 Homo sapiens cDNA
8888	21852		5.15	4.8E+00	AW750067.1	EST_HUMAN	PMO-BT0547-310100-002-504 BT0547 Homo sapiens cDNA
289	13384	26311	2.03	4.7E+00	BF240552.1	EST_HUMAN	601875654F1 NIH_MGC_55 Homo sapiens cDNA clone IMAGE:4090718 5'
280	13384	26311	1.69	4.7E+00	BF240552.1	EST_HUMAN	601875654F1 NIH_MGC_55 Homo sapiens cDNA clone IMAGE:4090718 5'
3287	16341	29280	1.68	4.7E+00	AL163280.2	NT	Homo sapiens chromosome 21 segment HS21C080
8084	21001	34397	0.55	4.6E+00	U87539.1	NT	Methanococcus jannaschii section 111 of 150 of the complete genome
9551	22513	35963	1.12	4.6E+00	BE648437.1	EST_HUMAN	7669g10.x1 NCI_CGAP_CLL1 Homo sapiens cDNA clone IMAGE:3292098 3' similar to TR:O75140 O75140
9551	22513	35964	1.12	4.6E+00	BE648437.1	EST_HUMAN	KIA00645 PROTEIN. contains element PTR5 repetitive element;
10755	23677		0.8	4.6E+00	AF240786.1	NT	KIA00645 PROTEIN. contains element PTR5 repetitive element;
8047	20884		0.61	4.5E+00	AF128177.1	NT	Homo sapiens glutathione S-transferase theta 2 (GSTT2) and glutathione S-transferase theta 1 (GSTT1)
11030	24811	36408	2.19	4.5E+00	AE001044.1	NT	genes, complete cds
12055	24928	39526	1.67	4.6E+00	BF689841.1	EST_HUMAN	Issatchenia orientalis inositolphosphoryltransferase synthase (IP-C1) gene, complete cds
3053	16110	28024	0.76	4.4E+00	BF530893.1	EST_HUMAN	Archaeoglobus fulgidus section 63 of 172 of the complete genome
3053	16110	28025	0.76	4.4E+00	BF530893.1	EST_HUMAN	602123238F1 NIH_MGC_56 Homo sapiens cDNA clone IMAGE:4280218 5'
6326	18396		1.69	4.4E+00	X13414.1	NT	602072585F1 NCI_CGAP_Bm67 Homo sapiens cDNA clone IMAGE:4215284 5'
6394	19462	32709	0.59	4.4E+00	AF156696.1	NT	602072585F1 NCI_CGAP_Bm67 Homo sapiens cDNA clone IMAGE:4215284 5'
6240	16313		0.71	4.3E+00	AF059679.1	NT	Murine I gene for MHC class II(a) associated invariant chain
7672	20630	33984	3.53	4.3E+00	Y13402.1	NT	Nicotiana tabacum inorganic phosphate transporter (PT1) mRNA, complete cds
7874	20818	34186	0.81	4.3E+00	AE001222.1	NT	Homo sapiens neutrophil collagenase (GLGNA) gene, promoter region and 5'UTR
11210	24164	37694	7.01	4.3E+00	AF240788.1	NT	Plasmodium falciparum R29R+var1 gene, exon 1
11279	24230		1.93	4.3E+00	11526311	NT	Treponema pallidum section 38 of 87 of the complete genome
5595	18691		3.57	4.2E+00	P16444	SWISSPROT	Homo sapiens glutathione S-transferase theta 2 (GSTT2) and glutathione S-transferase theta 1 (GSTT1)
5675	18770	31942	1.35	4.2E+00	P51826	SWISSPROT	genes, complete cds
5854	18945		0.58	4.2E+00	O27630	SWISSPROT	Homo sapiens DiGeorge syndrome critical region gene 2 (DGCR2), mRNA
6838	20162	33483	1.69	4.2E+00	P13983	SWISSPROT	MICROSOMAL DIFEPTIDASE PRECURSOR (MDP) (DEHYDROPEPTIDASE-1) (RENAL DIPEPTIDASE) (RDP)
6838	20162	33484	1.69	4.2E+00	P13983	SWISSPROT	LAF-4 PROTEIN (LYMPHOID NUCLEAR PROTEIN)
9311	22276	35708	5.1	4.2E+00	A1809013.1	EST_HUMAN	POTATIVE ATP-DEPENDENT HELICASE MTH1802
10278	22203	36688	1.1	4.2E+00	P31368	SWISSPROT	EXTENSIN PRECURSOR (CELL WALL HYDROXYPROLINE-RICH GLYCOPROTEIN)
10508	22430		0.53	4.2E+00	P40886	SWISSPROT	EXTENSIN PRECURSOR (CELL WALL HYDROXYPROLINE-RICH GLYCOPROTEIN)

Page 7 of 546
Table 4

Single Exon Probes Expressed in Bone Marrow

Probe SEQ ID NO:	Exon SEQ ID NO:	ORF SEQ ID NO:	Expression Signal	Most Similar (Top) Hit BLAST E Value	Top Hit Accession No.	Top Hit Database Source	Top Hit Descriptor
6046	26662	32334	0.64	4.1E+00	Q08185	SWISSPROT	CELLULAR TUMOR ANTIGEN P53
6046	26662	32335	0.64	4.1E+00	Q08185	SWISSPROT	CELLULAR TUMOR ANTIGEN P53
7319	20290	33633	0.76	4.1E+00	BE253868.1	EST_HUMAN	60110727F1 NIH_MGC_18 Homo sapiens cDNA clone IMAGE:3351634 5'
7420	20387	33738	0.65	4.1E+00	BF247830.1	EST_HUMAN	601859030F1 NIH_MGC_58 Homo sapiens cDNA clone IMAGE:4068758 5'
7923	20866	34254	7.87	4.1E+00	Q23810	SWISSPROT	YY1 PROTEIN PRECURSOR
8061	20868		0.64	4.1E+00	AB041523.1	NT	Patinopacten yessoensis mRNA for calcineurin A, complete cds
8066	21002	34398	3.95	4.1E+00	P28984	SWISSPROT	GENE 68 PROTEIN
8065	21002	34399	3.95	4.1E+00	P28984	SWISSPROT	GENE 68 PROTEIN
8249	21218	34628	2.5	4.1E+00	U57503.1	NT	Pan troglodytes novel repetitive solo LTR element in the RNU2 locus
9698	22861	36310	0.52	4.1E+00	P11253	SWISSPROT	50S RIBOSOMAL PROTEIN L4
10031	22858	36426	2.43	4.1E+00	BF692425.1	EST_HUMAN	602247838F1 NIH_MGC_62 Homo sapiens cDNA clone IMAGE:4333209 5'
10669	23591		0.45	4.1E+00	P46414	SWISSPROT	CYCLIN-DEPENDENT KINASE INHIBITOR 1B (CYCLIN-DEPENDENT KINASE INHIBITOR P27) (P27KIP1)
10981	23901	37414	0.47	4.1E+00	Q84242	SWISSPROT	3-OXOACYL-ACYL-CARRIER-PROTEIN] SYNTHASE III (BETA-KETOACYL-ACP SYNTHASE III) (KAS III)
11231	24184		2.3	4.1E+00	P09716	SWISSPROT	HYPOTHETICAL PROTEIN HVL1
11317	24267		13.22	4.1E+00	BE885880.1	EST_HUMAN	601507510F1 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:3903051 5'
3558	16602		0.7	4.0E+00	P38229	SWISSPROT	GLC7-INTERACTING PROTEIN 1
5534	20057	33361	0.94	4.0E+00	Q62653	SWISSPROT	SUCRASE-1SOMALTASE, INTESTINAL [CONTAINS: SUCRASE; ISOMALTASE]
5534	20057	33362	0.94	4.0E+00	Q62653	SWISSPROT	SUCRASE-1SOMALTASE, INTESTINAL [CONTAINS: SUCRASE; ISOMALTASE]
7123	20057	33361	1.01	4.0E+00	Q62653	SWISSPROT	SUCRASE-1SOMALTASE, INTESTINAL [CONTAINS: SUCRASE; ISOMALTASE]
7123	20057	33362	1.01	4.0E+00	Q62653	SWISSPROT	SUCRASE-1SOMALTASE, INTESTINAL [CONTAINS: SUCRASE; ISOMALTASE]
7395	20363	33715	1.47	4.0E+00	Q33010	SWISSPROT	CELL DIVISION PROTEIN FTSY HOMOLOG
8225	22191	35621	0.44	4.0E+00	Q14157	SWISSPROT	HYPOTHETICAL PROTEIN KIA0144
10303	23228	36711	0.43	4.0E+00	Q61309	SWISSPROT	NITRIC-OXIDE SYNTHASE (NOS, TYPE I) (NEURONAL NOS) (N-NOS) (NNOS)
10528	23448	36946	0.8	4.0E+00	AE002132.1	NT	Ureaplasma urealyticum section 33 of 58 of the complete genome
10620	23542	37041	0.49	4.0E+00	Q00511	SWISSPROT	URICASE (URATE OXIDASE)
10620	23542	37042	0.49	4.0E+00	Q00511	SWISSPROT	URICASE (URATE OXIDASE)
11802	23657	37461	1.67	4.0E+00	P14546	SWISSPROT	CYTOCHROME C OXIDASE POLYPEPTIDE III
11875	24757	36340	2.68	4.0E+00	P07564	SWISSPROT	GENOME POLYPROTEIN [CONTAINS: CAPSID PROTEIN C (CORE PROTEIN); MATRIX PROTEIN (ENVELOPE GLYCOPROTEIN M); MAJOR ENVELOPE PROTEIN E; NONSTRUCTURAL PROTEINS NS1, NS2A, NS2B, NS4A AND NS4B; HELICASE (NS3); RNA-DIRECTED RNA POLYMERASE (NS5)]

Page 8 of 546
Table 4
Single Exon Probes Expressed in Bone Marrow

Probe SEQ ID NO:	Exon SEQ ID NO:	ORF SEQ ID NO:	Expression Signal	Most Similar (Top) Hit BLAST E Value	Top Hit Accession No.	Top Hit Database Source	Top Hit Descriptor
11875	24757	38341	2.68	4.0E+00	P07584	SWISSPROT	GENOME POLYPROTEIN [CONTAINS: CAPSID PROTEIN C (CORE PROTEIN); MATRIX PROTEIN (ENVELOPE GLYCOPROTEIN M); MAJOR ENVELOPE PROTEIN E; NONSTRUCTURAL PROTEINS NS1, NS2A, NS2B, NS4A AND NS4B; HELICASE (NSS); RNA-DIRECTED RNA POLYMERASE (NS5)]
3513	18558	29483	4.66	3.9E+00	X64518.1	NT	N. tabacum chitinase gene 50 for class I chitinase C
4349	17378		0.89	3.9E+00	AF055468.1	NT	Mus musculus seminal vesicle secretory protein 89 (MSVSP89) gene, promoter region
5741	18835	32015	2.98	3.9E+00	BE814357.1	EST_HUMAN	MRO-BN0070-300500-028-H05 BN0070 Homo sapiens cDNA
5741	18835	32018	2.98	3.9E+00	BE814367.1	EST_HUMAN	MRO-BN0070-300500-028-H05 BN0070 Homo sapiens cDNA
6791	18845	33129	0.95	3.9E+00	AF288208.1	NT	Dicystotellum discoideum non-LTR retrotransposon TRE5-B, polyprotein (gag) and group-specific antigen (pol) genes, complete cds
6848	19801	33185	0.67	3.9E+00	U91328.1	NT	Human hereditary haemochromatosis region, histone 2A-like protein gene, hereditary haemochromatosis (HLA-H) gene, RoRet gene, and sodium phosphate transporter (NPT3) gene, complete cds
7057	20079	33388	4.24	3.9E+00	P38289	SWISSPROT	HYPOTHETICAL TRANSCRIPTIONAL REGULATOR IN AIDB-RPSF INTERGENIC REGION
7587	20548	33908	4.15	3.9E+00	M23907.1	NT	Human MHC class II lymphocyte antigen (DPw4-beta-1) gene, exon 2
8660	21628	35049	2.34	3.9E+00	X65885.1	NT	Xlaevis mRNA for M4 muscarinic receptor
11720	23617	37434	2.93	3.9E+00	Y18000.1	NT	Homo sapiens NF2 gene
11742	24827	38206	1.89	3.9E+00	AA681489.1	EST_HUMAN	nr18a12.s1 NCI_OGAP_Ew1 Homo sapiens cDNA clone IMAGE:1168318 similar to gb:A10416
2637	15836		1.27	3.9E+00	AE001682.1	NT	METALLOPROTEINASE INHIBITOR 1 PRECURSOR (HUMAN);
6630	19583	32854	0.93	3.9E+00	Q57830	SWISSPROT	Haecobacter pylori, strain J99 section 123 of the complete genome
6837	20161	33482	0.59	3.9E+00	A1493848.1	EST_HUMAN	HYPOTHETICAL PROTEIN MJ0385
8775	21742	35184	1.03	3.9E+00	D44725.1	EST_HUMAN	qz51807.x1 NCI_OGAP_Kd11 Homo sapiens cDNA clone IMAGE:2030437 5'
10164	23079		0.69	3.9E+00	A1390861.1	NT	HUMSUP135 Human brain cDNA Homo sapiens cDNA clone 148
12119	24989		15.21	3.9E+00	9831284	NT	Streptococcus cralis partial xpt gene for xanthine phosphoribosyltransferase, strain NCTC7864
4049	17088	28982	9.75	3.7E+00	AL161539.2	NT	Melanoplus sanguinipes entomopoxvirus, complete genome
7372	20342		0.9	3.7E+00	AL445065.1	NT	Arabidopsis thaliana DNA chromosome 4, contig fragment No. 39
8056	22022		0.49	3.7E+00	4503960	NT	Thermoplasma acidophilum complete genome, segment 3/5
9532	22485	35943	0.92	3.7E+00	U43541.1	NT	Homo sapiens glucokinase (hexokinase 4, maturity onset diabetes of the young 2) (GCK), nuclear gene encoding mitochondrial protein, mRNA
11760	24688	38269	1.73	3.7E+00	BF669279.1	EST_HUMAN	Mus musculus laminin beta 2 gene, exons 17-33, and complete cds
11760	24688	38270	1.73	3.7E+00	BF669279.1	EST_HUMAN	602120551F1 NIH_MGC_56 Homo sapiens cDNA clone IMAGE:4277748 5'
12256	25082		2.5	3.7E+00	AB013748.3	NT	602120551F1 NIH_MGC_56 Homo sapiens cDNA clone IMAGE:4277748 5'
565	13662	26575	4.04	3.6E+00	AV761055.1	EST_HUMAN	Gallus gallus mRNA for hypoxia-inducible factor-1 alpha, complete cds
							A761055 MDS Homo sapiens cDNA clone MDSBU10 5'

Page 9 of 546
Table 4

Single Exon Probes Expressed in Bone Marrow

Probe SEQ ID NO:	Exon SEQ ID NO:	ORF SEQ ID NO:	Expression Signal	Most Similar (Top) Hit BLAST E Value	Top Hit Accession No.	Top Hit Database Source	Top Hit Descriptor
5286	18292	31183	0.63	3.6E+00	Z89108.1	NT	<i>Bacillus subtilis</i> complete genome (section 6 of 21): from 889501 to 1208940
5327	18433	31185	0.73	3.6E+00	BF316316.1	EST_HUMAN	601801868F1 NIH_MGC_19 Homo sapiens cDNA clone IMAGE:4131016 5'
8897	21863	35286	0.93	3.6E+00	D12387.1	EST_HUMAN	HUM0001B08 Liver HepG2 cell line. Homo sapiens cDNA clone tb08
8897	21863	35286	0.93	3.6E+00	D12387.1	EST_HUMAN	HUM0001B08 Liver HepG2 cell line. Homo sapiens cDNA clone tb08
8992	21958	35383	4.21	3.6E+00	AE004447.1	NT	<i>Pseudomonas aeruginosa</i> PA01, section 8 of 529 of the complete genome
8992	21958	35384	4.21	3.6E+00	AE004447.1	NT	<i>Pseudomonas aeruginosa</i> PA01, section 8 of 529 of the complete genome
10022	22949	36416	0.44	3.6E+00	U72775.1	NT	<i>Ciconia episcopus</i> cytochrome b gene, mitochondrial gene encoding mitochondrial protein, partial cds
10022	22949	36417	0.44	3.6E+00	U72775.1	NT	<i>Ciconia episcopus</i> cytochrome b gene, mitochondrial gene encoding mitochondrial protein, partial cds
11200	24155		3.18	3.6E+00	M86795.1	NT	<i>Escherichia coli</i> glycero-phosphate dehydrogenase (gdpD) gene, partial cds; and the translation start site has been verified (gdpE), the translation start site has been verified (gdpG), and repressor protein (gdpR) genes, complete cds
6116	19183		1.1	3.5E+00	L42898.1	NT	<i>Borrelia burgdorferi</i> (strain 25015) outer surface protein (ospC) gene, partial cds
8337	19408	32847	0.96	3.5E+00	R18745.1	EST_HUMAN	Y44008.1 Soares Infant brain 1N1B Homo sapiens cDNA clone IMAGE:34940 5'
8087	21023	34422	0.6	3.6E+00	P07608	SWISSPROT	5-OXOPROLINASE (5-OXO-L-PROLINASE) (PYROGLUTAMASE) (5-OPASE)
8084	21030	34428	0.54	3.5E+00	AA892102.1	EST_HUMAN	cd37f10.61 Soares testis NHT Homo sapiens cDNA clone IMAGE:1618887 3' similar to gb.104213
8124	21061	34459	0.58	3.5E+00	4505284	NT	CELLULAR RETINALDEHYDE-BINDING PROTEIN (HUMAN);
8829	21796		0.6	3.5E+00	P24557	SWISSPROT	Homo sapiens macrophage stimulating 1 receptor (c-met-related tyrosine kinase) (MST1R) mRNA
8987	22352	35782	0.91	3.5E+00	AA190898.1	EST_HUMAN	THROMBOXANE-A SYNTHASE (TXA SYNTHASE) (TXS)
8987	22352	35783	0.91	3.5E+00	AA190898.1	EST_HUMAN	zq86804.61 Stratagene HeLa cell s3 837216 Homo sapiens cDNA clone IMAGE:827055 3' similar to
9850	22788	36240	0.96	3.6E+00	AL161553.2	NT	contains Alu repetitive element; contains element MSR1 repetitive element;
1514	14548	27517	5.3	3.4E+00	AF254577.1	NT	zq86804.61 Stratagene HeLa cell s3 837216 Homo sapiens cDNA clone IMAGE:827055 3' similar to
6883	19945	33241	0.49	3.4E+00	U77617.1	NT	contains Alu repetitive element; contains element MSR1 repetitive element;
7586	20547	33907	2.99	3.4E+00	P04052	SWISSPROT	<i>Arabidopsis thaliana</i> DNA chromosome 4, contig fragment No. 53
7868	20807	34288	0.9	3.4E+00	P04052	SWISSPROT	<i>Brassica napus</i> RPB5d mRNA, complete cds
9025	21891		0.68	3.4E+00	U65408.1	NT	<i>Chloranthus</i> yellow phytoplasmic acetate kinase gene, complete cds
9428	22392	35831	0.73	3.4E+00	AJ228042.1	NT	DNA-DIRECTED RNA POLYMERASE II LARGEST SUBUNIT
9487	22431	35889	0.55	3.4E+00	AJ250567.1	NT	DNA-DIRECTED RNA POLYMERASE II LARGEST SUBUNIT
10627	23549	37049	2.59	3.4E+00	AF013167.1	NT	Human alternatively spliced potassium channels ROM-K1, ROM-K2, ROM-K3, ROM-K4, ROM-K5, and ROM-K6 (KCNU1) gene, complete cds
							ROM-K6 (KCNU1) gene, complete cds
							Homo sapiens 850 kb contig between AML1 and CBR1 on chromosome 21q22, segment 2/3
							Homo sapiens partial TM6SF2 gene for tetraspanin protein, exon 8
							<i>Saccharomyces cerevisiae</i> MSS1 gene, complete cds

Table 4

Single Exon Probes Expressed In Bone Marrow

Probe SEQ ID NO:	Exon SEQ ID NO:	ORF SEQ ID NO:	Expression Signal	Most Similar (Top) Hit BLAST E Value	Top Hit Accession No.	Top Hit Database Source	Top Hit Descriptor
11861	24743	38327	2.88	3.4E+00	L7570.1	NT	Homo sapiens DiGeorge syndrome critical region, centromeric end
6188	18261	32498	1.03	3.3E+00	Q06689	SWISSPROT	PUTATIVE IRON ALCOHOL DEHYDROGENASE
6188	18261	32497	1.03	3.3E+00	Q06689	SWISSPROT	PUTATIVE IRON ALCOHOL DEHYDROGENASE
8225	21194	34601	0.9	3.3E+00	AF111168.2	NT	Homo sapiens serine palmitoyl transferase, subunit II gene, complete cds; and unknown genes
10830	23761	37250	0.9	3.3E+00	AP001511.1	NT	Bacillus halodurans genomic DNA, section 8/14
10830	23761	37251	0.9	3.3E+00	AP001511.1	NT	Bacillus halodurans genomic DNA, section 5/14
501	13573	28495	1.85	3.2E+00	X98422.1	NT	D.refo zp-50 POU gene
4056	13573	28495	0.88	3.2E+00	X98422.1	NT	D.refo zp-50 POU gene
4759	17779	30874	1.35	3.2E+00	4502404	NT	Homo sapiens carcinoembryonic antigen-related cell adhesion molecule 1 (biliary glycoprotein) (CEACAM1), mRNA
5648	18744	31910	1.2	3.2E+00	P54824	SWISSPROT	SQUALENE-HOPENE CYCLASE
5648	18744	31911	1.2	3.2E+00	P54824	SWISSPROT	SQUALENE-HOPENE CYCLASE
5683	18778	31950	2.79	3.2E+00	P12783	SWISSPROT	PHOSPHOGLYCERATE KINASE, CYTOSOLIC
5683	18778	31951	2.79	3.2E+00	P12783	SWISSPROT	PHOSPHOGLYCERATE KINASE, CYTOSOLIC
6439	18504	32754	1.66	3.2E+00	P18931	SWISSPROT	NADH-UBIQUINONE OXIDOREDUCTASE CHAIN 4
6439	18504	32755	1.66	3.2E+00	P18931	SWISSPROT	NADH-UBIQUINONE OXIDOREDUCTASE CHAIN 4
7863	20807	34185	0.71	3.2E+00	P04275	SWISSPROT	VON WILLEBRAND FACTOR PRECURSOR (VWF)
8052	20889	34388	2.47	3.2E+00	Y13655.1	NT	Chlamydomonas reinhardtii chloroplast DNA for rps8, ycf4, ycf3, rps18 genes
8052	20889	34387	2.47	3.2E+00	Y13655.1	NT	Chlamydomonas reinhardtii chloroplast DNA for rps8, ycf4, ycf3, rps18 genes
8385	22350		4.78	3.2E+00	P13061	SWISSPROT	PERIPLASMIC [NIFE] HYDROGENASE SMALL SUBUNIT (NIFE HYDROGENLYASE SMALL CHAIN)
8888	22841	36298	1.87	3.2E+00	M36383.1	NT	S.cerevisiae trehalase deaminase (TLV1) gene, complete cds
10500	28422	36821	2	3.2E+00	AB016081.2	NT	Oryzias latipes OIGC8 gene for glutaryl cyclase C, complete cds
12217	26055		2.84	3.2E+00	L33836.1	NT	Sus scrofa choline acetyltransferase gene, promoter region
5875	18080	32281	2.25	3.1E+00	Q10135	SWISSPROT	HYPOTHETICAL 142.5 KD PROTEIN C23E2.02 IN CHROMOSOME 1
7618	20578	33941	0.9	3.1E+00	P52178	SWISSPROT	TRIOSE PHOSPHATE/PHOSPHATE TRANSLOCATOR, NON-GREEN PLASTID PRECURSOR (CTP1)
7889	20838		1	3.1E+00	AF303225.1	NT	Bacillus alcalophilus peckate lyase (pelE) gene, complete cds
8424	21383	34804	0.43	3.1E+00	P40985	SWISSPROT	PROBABLE UBIQUITIN-PROTEIN LIGASE HUL4
8849	21915	35338	4.37	3.1E+00	P48894	SWISSPROT	TYPE I IODOETHYRINE DEIODINASE (TYPE I 5DEIODINASE) (DIOI) (TYPE 1 DII) (5DII)
8849	21915	35339	4.37	3.1E+00	P48894	SWISSPROT	TYPE I IODOETHYRINE DEIODINASE (TYPE I 5DEIODINASE) (DIOI) (TYPE 1 DII) (5DII)
							GLUTAMATE [NMDA] RECEPTOR SUBUNIT EPSILON 3 PRECURSOR (N-METHYL D-ASPARTATE RECEPTOR SUBTYPE 2C) (NR2C) (NMDAR2C)
9814	22618		3.9	3.1E+00	Q14957	SWISSPROT	COLLAGEN ALPHA 2(I) CHAIN PRECURSOR
9880	22883	36089	0.48	3.1E+00	Q01149	SWISSPROT	Chlorella vulgaris chloroplast, complete genome
10258	23181	36688	0.86	3.1E+00	7524759	NT	
10347	23271		0.61	3.1E+00	Q10125	SWISSPROT	HYPOTHETICAL 58.3 KD PROTEIN F52C8.5 IN CHROMOSOME III

Page 11 of 546
Table 4

Single Exon Probes Expressed in Bone Marrow

Probe SEQ ID NO:	Exon SEQ ID NO:	ORF SEQ ID NO:	Expression Signal	Most Similar (Top) Hit BLAST E Value	Top Hit Accession No.	Top Hit Database Source	Top Hit Descriptor
10700	23622	37118	5.2	3.1E+00	P49365	SWISSPROT	DEOXYHYPUSINE SYNTHASE (DHS)
11791	23846		1.98	3.1E+00	P33515	SWISSPROT	GENOME POLYPROTEIN [CONTAINS: CAPSID PROTEIN C (CORE PROTEIN); MATRIX PROTEIN (ENVELOPE PROTEIN M); MAJOR ENVELOPE PROTEIN E; NONSTRUCTURAL PROTEINS NS1, NS2A, NS2B, NS4A AND NS4B; HELICASE (NS3); RNA-DIRECTED RNA POLYMERASE (NS6)]
11811	24698		2.78	3.1E+00	S68880.1	NT	retinoid acid nuclear receptor isoform beta 2 [mice, embryonal carcinoma cell line, POC7-MZ1, mRNA, 2871 nt]
2849	15909	28833	1.5	3.0E+00	8923884	NT	Homo sapiens hypothetical protein PRO0889 (PRO0889), mRNA
5411	18514	31392	1.33	3.0E+00	X53098.1	NT	S. aureus genes encoding Sau661 DNA methyltransferase and Sau661 restriction endonuclease
6708	19784	33043	0.79	3.0E+00	X56037.1	NT	Corynebacterium glutamicum thrC gene for threonine synthase (EC 4.2.99.2)
6708	19784	33044	0.79	3.0E+00	X56037.1	NT	Corynebacterium glutamicum thrC gene for threonine synthase (EC 4.2.99.2)
7363	20333		9.68	3.0E+00	P18406	SWISSPROT	CYR61 PROTEIN PRECURSOR (3CH61)
7402	20370		0.59	3.0E+00	Q13201	SWISSPROT	ENDOTHELIAL CELL MULTIMERIN PRECURSOR
8258	22224		1.21	3.0E+00	X67838.1	NT	B. napus DNA for myrosinase
10657	23679	37076	0.54	3.0E+00	Q58805	SWISSPROT	S-ADENOSYLMETHIONINE SYNTHETASE (METHIONINE ADENOSYLTRANSFERASE) (ADOMET SYNTHETASE)
11008	23973	37497	1.65	3.0E+00	Q16181	SWISSPROT	CDC10 PROTEIN HOMOLOG
11351	24301	37827	4.64	3.0E+00	P51842	SWISSPROT	RETINAL GUANYLYL CYCLASE 2 PRECURSOR (GUANYLATE CYCLASE 2F, RETINAL) (RETGC-2) (ROD OUTER SEGMENT MEMBRANE GUANYLATE CYCLASE 2) (ROS-GC2) (GUANYLATE CYCLASE F) (GC-F)
11351	24301	37828	4.64	3.0E+00	P51842	SWISSPROT	RETINAL GUANYLYL CYCLASE 2 PRECURSOR (GUANYLATE CYCLASE 2F, RETINAL) (RETGC-2) (ROD OUTER SEGMENT MEMBRANE GUANYLATE CYCLASE 2) (ROS-GC2) (GUANYLATE CYCLASE F) (GC-F)
2028	15048	28059	2.33	2.9E+00	AE002225.2	NT	Chlamydia pneumoniae AR39, section 53 of 94 of the complete genome
6182	19288		0.8	2.9E+00	AB026033.1	NT	Bonapartha pediculus mitochondrial DNA for 16S ribosomal RNA
7094	20028	33332	1.97	2.9E+00	Z36879.1	NT	F. pringlei gdcSP4 gene for P-protein of the glycine cleavage system
7418	20385	33734	5.15	2.9E+00	O14514	SWISSPROT	BRAIN-SPECIFIC ANGIOGENESIS INHIBITOR 1 PRECURSOR
7418	20385	33735	5.15	2.9E+00	O14514	SWISSPROT	BRAIN-SPECIFIC ANGIOGENESIS INHIBITOR 1 PRECURSOR
7689	20647	34011	5.32	2.9E+00	P46589	SWISSPROT	ADHERENCE FACTOR (ADHESION AND AGGREGATION MEDIATING SURFACE ANTIGEN)
8200	21170	34580	0.61	2.9E+00	P05844	SWISSPROT	STRUCTURAL POLYPROTEIN [CONTAINS: MAJOR STRUCTURAL PROTEIN VP2; MINOR STRUCTURAL PROTEIN VP3]
8200	21170	34581	0.61	2.9E+00	P05844	SWISSPROT	STRUCTURAL POLYPROTEIN [CONTAINS: MAJOR STRUCTURAL PROTEIN VP2; MINOR STRUCTURAL PROTEIN VP3]
8434	21403	34818	0.82	2.9E+00	BF344171.1	EST_HUMAN	NONSTRUCTURAL PROTEIN VP4; MINOR STRUCTURAL PROTEIN VP3
1451	14484	27460	4.16	2.8E+00	AF186398.1	NT	602017413F1 NCL_OGAP_Bmr64 Homo sapiens cDNA clone IMAGE:4163069 5'
							Buxus harlandii maturase K (matK) gene, partial cds; chloroplast gene for chloroplast product

Page 12 of 546
Table 4

Single Exon Probes Expressed In Bone Marrow

Probe SEQ ID NO:	Exon SEQ ID NO:	ORF SEQ ID NO:	Expression Signal	Most Similar (Top) Hit BLAST E Value	Top Hit Accession No.	Top Hit Database Source	Top Hit Descriptor
1637	14689		2.57	2.8E+00	AL161552.2	NT	Arabidopsis thaliana DNA chromosome 4, contig fragment No. 52
7529	20492	33854	4.93	2.8E+00	8393724	NT	Mus musculus endomucin (LOC53423), mRNA
6972	22899		0.58	2.8E+00	BE566182.1	EST_HUMAN	601342768F1 NIH_MGC 53 Homo sapiens cDNA clone IMAGE:3684807 5'
11048	20492	33854	1.73	2.8E+00	8393724	NT	Mus musculus endomucin (LOC53423), mRNA
233	19333	28258	13.38	2.7E+00	6679308	NT	Mus musculus per-hexamer repeat gene 3 (Phx3), mRNA
233	19333	28257	13.38	2.7E+00	6679308	NT	Mus musculus per-hexamer repeat gene 3 (Phx3), mRNA
5631	18727	31888	1.11	2.7E+00	L14005.1	NT	Homo sapiens apoa polymorphism Kingle IV gene, exons 1 and 2
8485	21453		0.68	2.7E+00	U15947.1	NT	Ipomoea purpurea chalcone synthase (CHS) gene including complete 5'UTR and complete cds
9319	22284		1.69	2.7E+00	AL116459.1	NT	Borhytia chereia strain T4 cDNA library under conditions of nitrogen deprivation
9787	21110	34510	0.64	2.7E+00	AW088191.1	EST_HUMAN	xc88612.x1 NCI_CGAP_Birc95 Homo sapiens cDNA clone IMAGE:2591374 3' similar to gb:M17753
10896	23788		1.69	2.7E+00	BE069327.1	EST_HUMAN	THYMOSIN BETA-4 (HUMAN);
4701	17722	30615	6.29	2.6E+00	AF068749.1	NT	CMO-BT0281-031799-087-H04 BT0281 Homo sapiens cDNA
5827	18723	31883	2.08	2.6E+00	6755601	NT	Mus musculus sphingosine kinase (SPHK1b) mRNA, complete cds
5827	18723	31884	2.08	2.6E+00	6755601	NT	Mus musculus SRY-box containing gene 13 (Sox13), mRNA
5925	19011		3.9	2.6E+00	Y17082.1	NT	Mus musculus SRY-box containing gene 13 (Sox13), mRNA
7803	26002		0.7	2.6E+00	AJ224639.1	NT	Mycobacterium fortuitum furA II gene
7807	20906		32.15	2.6E+00	AF235502.1	NT	Homo sapiens Surf-5 and Surf-6 genes
8394	21363	34770	1.12	2.6E+00	AJ132180.1	NT	Mus musculus SH2-containing inositol 5-phosphatase (Ship) gene, exons 16 through 27, and complete cds
8394	21363	34771	1.12	2.6E+00	AJ132180.1	NT	faba bean necrotic yellows virus C2-Eg gene, isolate Egyptian EV1-83
10015	22942	35408	3.12	2.6E+00	AL161540.2	NT	faba bean necrotic yellows virus C2-Eg gene, isolate Egyptian EV1-83
10720	23642		1.61	2.6E+00	8055193	NT	Arabidopsis thaliana DNA chromosome 4, contig fragment No. 40
12841	25977		2.58	2.6E+00	11419220	NT	Mus musculus cleavage and polyadenylation specificity factor 3 (Cpsf3), mRNA
1460	14483	27466	3.73	2.6E+00	AJ271844.1	NT	Homo sapiens ATP-binding cassette, sub-family B (MDR/TAP), member 4 (ABCB4), mRNA
1460	14493	27467	3.73	2.6E+00	AJ271844.1	NT	Aspergillus nidulans recQ gene for DNA helicase, exons 1-4
5911	18997	32186	2.32	2.6E+00	P13485	SW/ISSPROT	Aspergillus nidulans recQ gene for DNA helicase, exons 1-4
5911	18997	32187	2.32	2.6E+00	P13485	SW/ISSPROT	TEICHOIC ACID BIOSYNTHESIS PROTEIN F
6801	18997	32186	1.49	2.6E+00	P13485	SW/ISSPROT	TEICHOIC ACID BIOSYNTHESIS PROTEIN F
6801	18997	32187	1.49	2.6E+00	P13485	SW/ISSPROT	TEICHOIC ACID BIOSYNTHESIS PROTEIN F
6892	18944	33240	0.66	2.6E+00	D30052.1	NT	TEICHOIC ACID BIOSYNTHESIS PROTEIN F
7960	20901	34293	0.55	2.6E+00	P17598	SW/ISSPROT	Vibrio cholerae ctaA gene and ctaB gene for cholera toxin, complete cds
8035	20972	34366	0.97	2.6E+00	AW949158.1	EST_HUMAN	LATENCY-RELATED PROTEIN 1
8095	21031	34429	0.51	2.6E+00	4502902	NT	QV4-F10005-110500-205-g07 FT0005 Homo sapiens cDNA
							Homo sapiens clathrin, heavy polypeptide-like 1 (CLTCL1) mRNA

Page 13 of 546
Table 4

Single Exon Probes Expressed in Bone Marrow

Probe SEQ ID NO:	Exon SEQ ID NO:	ORF SEQ ID NO:	Expression Signal	Most Similar (Top) Hit BLAST E Value	Top Hit Accession No.	Top Hit Database Source	Top Hit Descriptor
9459	22423	35861	1.49	2.9E+00	D50307.1	NT	Rice DNA for aldolase C-1, complete cds
10211	23136	36623	0.73	2.9E+00	BE297768.1	EST_HUMAN	801175779F1 NIH_MGC_17 Homo sapiens cDNA clone IMAGE:3531090 5'
12214	25053		2.5	2.9E+00	AF289685.1	NT	Mus musculus EIF-4H gene, partial cds; LIMK1 gene, complete cds; and ELN gene, partial cds
3026	16083	28008	1.1	2.4E+00	M24282.1	NT	Chicken alpha-3 collagen type VI mRNA, 3' and
4944	17980	30851	5.31	2.4E+00	4603352	NT	Homo sapiens double C2-like domains, alpha (DOC2A) mRNA
6126	18204	32427	3.62	2.4E+00	P02843	SWISSPROT	VITELLOGENIN I PRECURSOR (YOLK PROTEIN 1)
7606	20567	33927	0.71	2.4E+00	BF667502.1	EST_HUMAN	802120856F1 NIH_MGC_58 Homo sapiens cDNA clone IMAGE:4278012 5'
7606	20567	33928	0.71	2.4E+00	BF667502.1	EST_HUMAN	802120856F1 NIH_MGC_58 Homo sapiens cDNA clone IMAGE:4278012 5'
8479	21448	34864	2.08	2.4E+00	P28842	SWISSPROT	CD27L RECEPTOR PRECURSOR (T-CELL ACTIVATION ANTIGEN CD27) (T14)
8479	21448	34865	2.08	2.4E+00	P28842	SWISSPROT	CD27L RECEPTOR PRECURSOR (T-CELL ACTIVATION ANTIGEN CD27) (T14)
8554	21522		2.92	2.4E+00	AEO01486.1	NT	Helicobacter pylori, strain J99 section 47 of 132 of the complete genome
8698	21984		1.46	2.4E+00	AW875126.1	EST_HUMAN	RC2-PT0004-031289-011-405 PT0004 Homo sapiens cDNA
9180	22146	35573	9.52	2.4E+00	P24091	SWISSPROT	ENDOCHITINASE B PRECURSOR (CHIN-B)
10398	23320	36803	2.34	2.4E+00	P13673	SWISSPROT	SKIN GRANULE PROTEIN PRECURSOR
10398	23320	36804	2.34	2.4E+00	P13673	SWISSPROT	SKIN GRANULE PROTEIN PRECURSOR
10466	23388	36882	2.1	2.4E+00	X92511.1	NT	H. sapiens CTGF gene and promoter region
10604	23528		6.1	2.4E+00	P09089	SWISSPROT	XYLULOSE KINASE (XYLULOKINASE)
10685	23607	37100	1.67	2.4E+00	BE326702.1	EST_HUMAN	hr63f06.x1 NC1 CGAP_Kd11 Homo sapiens cDNA clone IMAGE:3133187 3'
10685	23607	37101	1.67	2.4E+00	BE326702.1	EST_HUMAN	hr63f06.x1 NC1 CGAP_Kd11 Homo sapiens cDNA clone IMAGE:3133187 3'
10655	23875	37389	1.14	2.4E+00	Q51481	SWISSPROT	DENTRIFICATION REGULATORY PROTEIN N1RQ
11415	24359	37694	2.2	2.4E+00	Y14079.1	NT	Bacillus subtilis chromosomal DNA, region 75 degrees: gipPFKD operon and downstream
11517	24458		1.66	2.4E+00	AF066872.1	NT	Capra hircus alphaS2-casein type C gene, intron 15
11686	24652	38231	2.14	2.4E+00	AF169682.2	NT	Fragaria x anemessa cytosolic ascorbate peroxidase (ApxSC) gene, ApxSC-c allele, complete cds
1258	14293	27257	11.33	2.3E+00	Z46724.1	NT	G.domesticus artificial single chain antibody gene (L3)
4147	17178		1.45	2.3E+00	AL401081.1	NT	Bos taurus partial cytb gene for cytochrome b
5834	19020		0.89	2.3E+00	N86245.1	EST_HUMAN	J7340F Human fetal heart, Lambda ZAP Express Homo sapiens cDNA clone J7340 5' similar to
7687	20645	34009	2.42	2.3E+00	0678554	NT	PROLYLCARBOXYPEPTIDASE
7853	25003		2.79	2.3E+00	P07199	SWISSPROT	Rattus norvegicus ATPase, Cat+ transporting, ubiquitous (Atp2a3), mRNA
8059	20906	34392	1.12	2.3E+00	X60265.1	NT	MAJOR CENTROMERE AUTOANTIGEN B (CENTROMERE PROTEIN B) (CENP-B)
9465	22428	35868	0.54	2.3E+00	5835317	NT	M.mazal dnaK and dnaJ genes homologues coding for DnaK and DnaJ
							Polyporus ornithinis mitochondrion, complete genome
							ALPHA(1,3)-FUCOSYLTRANSFERASE (GALACTOSIDE 3-L-FUCOSYLTRANSFERASE)
9525	22488	35936	2.03	2.3E+00	Q11127	SWISSPROT	(FUCOSYLTRANSFERASE 4) (FUCT-IV)

Page 14 of 546
Table 4

Single Exon Probes Expressed in Bone Marrow

Probe SEQ ID NO:	Exon SEQ ID NO:	ORF SEQ ID NO:	Expression Signal	Most Similar (Top) Hit BLAST E Value	Top Hit Accession No.	Top Hit Database Source	Top Hit Descriptor
10858	23878	37390	0.45	2.3E+00	P02461	SWISSPROT	COLLAGEN ALPHA 1(III) CHAIN PRECURSOR
11153	24113	37638	7.84	2.3E+00	Q07078	SWISSPROT	ANNEXIN VII (SYNEXIN)
11887	24778	38364	1.59	2.3E+00	P45831	SWISSPROT	HYPOTHETICAL 171.0 KD PROTEIN IN SPOIIIC-CWLA INTERGENIC REGION
12072	24945	38539	2.34	2.3E+00	BF541987.1	EST_HUMAN	602069121F1 NIH_MGC_58 Homo sapiens cDNA clone IMAGE:4068173 5'
12072	24945	38540	2.34	2.3E+00	BF541987.1	EST_HUMAN	602069121F1 NIH_MGC_58 Homo sapiens cDNA clone IMAGE:4068173 5'
12443	25205	31628	6.31	2.3E+00	BE885237.1	EST_HUMAN	601433673F1 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:3918843 5'
13077	25609		1.3	2.3E+00	AF281862.1	NT	Neurospora crassa G protein alpha subunit GNA-3 (gna-3) gene, complete cds
4046	17084	29981	1.42	2.2E+00	AF020528.1	NT	Magnaporthe grisea Class IV chitin synthase (chs4) gene, complete cds
4342	17369	30251	3.8	2.2E+00	D67071.1	NT	Rat gene for regucalcin, exon1 (non-coding exon)
4342	17369	30252	3.8	2.2E+00	D67071.1	NT	Rat gene for regucalcin, exon1 (non-coding exon)
5415	18518	31396	11.02	2.2E+00	O88307	SWISSPROT	SORTILIN-RELATED RECEPTOR PRECURSOR (SORTING PROTEIN-RELATED RECEPTOR CONTAINING LDLR CLASS A REPEATS) (MSORLA) (SORLA-1) (LOW-DENSITY LIPOPROTEIN RECEPTOR RELATIVE WITH 11 LIGAND-BINDING REPEATS) (LDLR RELATIVE WITH 11 LIGAND-BINDING REPEATS) (LR11) (<
5415	18518	31396	11.02	2.2E+00	O88307	SWISSPROT	SORTILIN-RELATED RECEPTOR PRECURSOR (SORTING PROTEIN-RELATED RECEPTOR CONTAINING LDLR CLASS A REPEATS) (MSORLA) (SORLA-1) (LOW-DENSITY LIPOPROTEIN RECEPTOR RELATIVE WITH 11 LIGAND-BINDING REPEATS) (LDLR RELATIVE WITH 11 LIGAND-BINDING REPEATS) (LR11) (<
5953	19038	32234	1.83	2.2E+00	BE927220.1	EST_HUMAN	RC3-CT0254-300800-022-c08 CT0254 Homo sapiens cDNA
5953	19038	32235	1.83	2.2E+00	BE927220.1	EST_HUMAN	RC3-CT0254-300800-022-c08 CT0254 Homo sapiens cDNA
6180	19255	32488	8.78	2.2E+00	BE260383.1	EST_HUMAN	600843401T1 NIH_MGC_17 Homo sapiens cDNA clone IMAGE:2958777 3'
6489	19594	32804	3.91	2.2E+00	Q00335	SWISSPROT	MINOR VIRION STRUCTURAL PROTEIN MU-2
6750	19804	33085	3.14	2.2E+00	P51459	SWISSPROT	INSULIN-LIKE GROWTH FACTOR II PRECURSOR (IGF-II) (SOMATOMEDIN A)
7149	18381		3.5	2.2E+00	AA594574.1	EST_HUMAN	m95802.s1 NCI_CGAP_Co10 Homo sapiens cDNA clone IMAGE:1058379 3'
7839	20502	33881	0.95	2.2E+00	AA137027.1	EST_HUMAN	zn8704.J1 Stratiogene fetal retina 937202 Homo sapiens cDNA clone IMAGE:568143 5'
7865	20809	34187	18.24	2.2E+00	AA449012.1	EST_HUMAN	zn05g10.J1 Soares fetal fetus_Nb2HF9_9w Homo sapiens cDNA clone IMAGE:786634 5'
7953	20894	34287	0.98	2.2E+00	P54918	SWISSPROT	ALANINE RACEMASE
8439	21408	34820	0.69	2.2E+00	BE301560.1	EST_HUMAN	bb17h12.x1 NIH_MGC_21 Homo sapiens cDNA clone IMAGE:2963207 3' similar to gb:D45836 Mouse mRNA for nuclear pore-targeting-complex component of (MOUSE);
8439	21408	34821	0.69	2.2E+00	BE301560.1	EST_HUMAN	bb17h12.x1 NIH_MGC_21 Homo sapiens cDNA clone IMAGE:2963207 3' similar to gb:D45836 Mouse mRNA for nuclear pore-targeting-complex component of (MOUSE);
9887	22850		11.22	2.2E+00	BE741678.1	EST_HUMAN	601594733F1 NIH_MGC_9 Homo sapiens cDNA clone IMAGE:3948561 5'
9927	25698		2.53	2.2E+00	Q04706	SWISSPROT	TRANSPONSON TY1 PROTEIN A

Page 15 of 546
Table 4

Single Exon Probes Expressed in Bone Marrow

Probe SEQ ID NO:	Exon SEQ ID NO:	ORF SEQ ID NO:	Expression Signal	Most Similar (Top) Hit BLAST E Value	Top Hit Accession No.	Top Hit Database Source	Top Hit Descriptor
10413	23335	36819	1.57	2.2E+00	AI280373.1	EST_HUMAN	qin69b03.x1 Soares placenta_81c9weeks_2N1bIP81c9W Homo sapiens cDNA clone IMAGE:1893965 3' similar to gb:Y00433 GLUTATHIONE PEROXIDASE (HUMAN);
10413	23335	36820	1.57	2.2E+00	AI280373.1	EST_HUMAN	qin69b03.x1 Soares placenta_81c9weeks_2N1bIP81c9W Homo sapiens cDNA clone IMAGE:1893965 3' similar to gb:Y00433 GLUTATHIONE PEROXIDASE (HUMAN);
10458	23378	36871	2.22	2.2E+00	BF246782.1	EST_HUMAN	601855591F1 NIH_MGC_57 Homo sapiens cDNA clone IMAGE:4076391 5'
10822	23743	37244	3.06	2.2E+00	AF183416.1	NT	Homo sapiens ovarian granulosa cell 13.0 kDa protein hGR74 homolog mRNA, complete cds
11768	23923	37442	3.23	2.2E+00	P07911	SWISSPROT	UROMODULIN PRECURSOR (TAMM-HORSFALL URINARY GLYCOPROTEIN) (THP)
11937	24818	38415	8.31	2.2E+00	P10407	SWISSPROT	EARLY E1A 28 KD PROTEIN
571	15844	26555	12.39	2.1E+00	AF132812.2	NT	Mus musculus pre-T cell receptor alpha gene, enhancer region and upstream region
3601	16646		0.83	2.1E+00	AW449368.1	EST_HUMAN	UIH-B13-akt-e-08-0-UI.s1 NCI CGAP Sub5 Homo sapiens cDNA clone IMAGE:2734550 3'
6255	18326		0.86	2.1E+00	P76357	SWISSPROT	HYPOTHETICAL PROTEIN MG302 HOMOLOG
6980	20203	33532	3.45	2.1E+00	O70159	SWISSPROT	ALPHA-2-HS-GLYCOPROTEIN PRECURSOR (FETUIN-A)
							Homo sapiens dystralin, limb girdle muscular dystrophy 2B (autosomal recessive) (DYSLF) mRNA, and translated products
7225	20247	33581	0.61	2.1E+00	4503430	NT	
							yy08a10.s1 Soares melanocyte 2N1bIM Homo sapiens cDNA clone IMAGE:270818 3' similar to gb:M55064 TRANSCRIPTION INITIATION FACTOR TFIIID (HUMAN);
7246	19981	33278	5.97	2.1E+00	N28575.1	EST_HUMAN	AU123630 NT2RM2 Homo sapiens cDNA clone NT2RM2000071 5'
8842	21809		1.82	2.1E+00	AU123630.1	EST_HUMAN	Homo sapiens p220del (DOKDEL) mRNA, complete cds
1201	14240	27197	1.39	2.0E+00	AF180527.1	NT	Homo sapiens p220del (DOKDEL) mRNA, complete cds
1201	14240	27198	1.39	2.0E+00	AF180527.1	NT	Oryctolagus cuniculus Nerv.K+-ATPase beta 1 subunit mRNA, complete cds
1338	14372	27342	1.19	2.0E+00	AF204927.1	NT	PUTATIVE RNA METHYLTRANSFERASE SPB1
1578	14611		3.42	2.0E+00	P25582	SWISSPROT	R.norvegicus mRNA for collagen alpha1 type I
2159	15175	28105	4.98	2.0E+00	Z78278.1	NT	R.norvegicus mRNA for collagen alpha1 type I
2159	15175	28106	4.98	2.0E+00	Z78278.1	NT	M13c05.x1 NCI CGAP GU1 Homo sapiens cDNA clone IMAGE:2872168 3' similar to gb:X01677 GLYCERALDEHYDE 3-PHOSPHATE DEHYDROGENASE, LIVER (HUMAN);
4127	17160	30048	2.24	2.0E+00	AW684496.1	EST_HUMAN	M13c05.x1 NCI CGAP GU1 Homo sapiens cDNA clone IMAGE:2872168 3' similar to gb:X01677 GLYCERALDEHYDE 3-PHOSPHATE DEHYDROGENASE, LIVER (HUMAN);
4127	17160	30049	2.24	2.0E+00	AW684496.1	EST_HUMAN	STRUCTURAL POLYPEPTIDE [CONTAINS: NUCLEOCAPSID PROTEIN C; MEMBRANE GLYCOPROTEINS E1 AND E2]
7798	20750		0.85	2.0E+00	P07566	SWISSPROT	Escherichia coli 0157 DNA, map position at 46 min., complete cds
8358	21327	34738	3.84	2.0E+00	AB008678.1	NT	Escherichia coli 0157 DNA, map position at 46 min., complete cds
8358	21327	34739	3.84	2.0E+00	AB008678.1	NT	Escherichia coli 0157 DNA, map position at 46 min., complete cds
8358	21327	34740	3.84	2.0E+00	AB008678.1	NT	Escherichia coli 0157 DNA, map position at 46 min., complete cds
8274	22240	35688	3.22	2.0E+00	F31500.1	EST_HUMAN	HSPD22703 H3M3 Homo sapiens cDNA clone s4000117B08
12756	25638	31427	4.52	2.0E+00	5834943	NT	Gallus gallus mitochondrion, complete genome

Table 4

Single Exon Probes Expressed in Bone Marrow

Probe SEQ ID NO:	Exon SEQ ID NO:	ORF SEQ ID NO:	Expression Signal	Most Similar (Top) Hit BLAST E Value	Top Hit Accession No.	Top Hit Database Source	Top Hit Descriptor
4829	17846	30746	0.98	1.9E+00	AF209468.1	NT	Danio rerio Rth50-like protein mRNA, complete cds
5679	18774	31945	4.52	1.9E+00	6754389	NT	Mus musculus Inositol 1,4,5-triphosphate receptor 1 (Itpr1), mRNA
5679	18774	31946	4.52	1.9E+00	6754389	NT	Mus musculus Inositol 1,4,5-triphosphate receptor 1 (Itpr1), mRNA
6219	19293	32527	1.05	1.9E+00	BE969695.1	EST_HUMAN	601679639F1 NIH_MGC_78 Homo sapiens cDNA clone IMAGE:3946881 5'
6810	19884		1.19	1.9E+00	AW845889.1	EST_HUMAN	MRO-CT0063-071038-002-g02 CT0063 Homo sapiens cDNA
6912	19884		2.37	1.9E+00	Q63627	SWISSPROT	GTD-BINDING SR-LIKE PROTEIN RA4
8802	21769	35193	1.72	1.9E+00	P02467	SWISSPROT	COLLAGEN ALPHA 2(I) CHAIN PRECURSOR
8802	21769	35194	1.72	1.9E+00	P02467	SWISSPROT	COLLAGEN ALPHA 2(I) CHAIN PRECURSOR
9006	21972		3.6	1.9E+00	BF360206.1	EST_HUMAN	CM3-MT0114-010800-323-412 MT0114 Homo sapiens cDNA
9245	22211		1.52	1.9E+00	O51781	SWISSPROT	ARGININE DEIMINASE (ADI) (ARGININE DIHYDROLASE) (AD)
9868	22915	36380	0.53	1.9E+00	AA669126.1	EST_HUMAN	ab54904.s1 Stratagene lung (#8372710) Homo sapiens cDNA clone IMAGE:854574 3' similar to contains Alu repetitive element; contains element L1 L1 repetitive element;
10829	23849	37364	0.63	1.9E+00	AF248289.1	NT	Homo sapiens gag-pro-pol precursor protein gene, partial cds
3109	16166	29077	1.5	1.8E+00	P21004	SWISSPROT	PROTEIN B8 PRECURSOR
3137	16194	29103	10.81	1.8E+00	U04353.1	NT	Synechococcus sp. PCC7942 copper transporting P-ATPase (ctaA) and ATP synthase epsilon subunit (atpE) genes, complete cds
3137	16194	29104	10.81	1.8E+00	U04353.1	NT	Synechococcus sp. PCC7942 copper transporting P-ATPase (ctaA) and ATP synthase epsilon subunit (atpE) genes, complete cds
5967	19052		1.84	1.8E+00	P18502	SWISSPROT	HEDGEHOG RECEPTOR (PATCHED PROTEIN)
6224	19288	32532	2.2	1.8E+00	BF311899.1	EST_HUMAN	601897854F1 NIH_MGC_19 Homo sapiens cDNA clone IMAGE:4127384 5'
6538	19600		1.19	1.8E+00	BF683327.1	EST_HUMAN	602139470F1 NIH_MGC_46 Homo sapiens cDNA clone IMAGE:4298272 5'
6904	19666	33253	1.02	1.8E+00	BF305652.1	EST_HUMAN	601893489F1 NIH_MGC_17 Homo sapiens cDNA clone IMAGE:4139038 5'
7260	19995	33262	1.07	1.8E+00	P21249	SWISSPROT	MAJOR ANTIGEN
7476	20442		0.7	1.8E+00	P27127	SWISSPROT	LIPOPOLYSACCHARIDE 1,6-GALACTOSYLTRANSFERASE (UDP-D-GALACTOSE--(GLUCOSYL)LIPOPOLYSACCHARIDE-ALPHA-1,3-D-GALACTOSYLTRANSFERASE)
8454	21423	34838	0.9	1.8E+00	P11369	SWISSPROT	RETROVIRUS-RELATED POL POLYPROTEIN [CONTAINS: REVERSE TRANSCRIPTASE; ENDONUCLEASE]
8454	21423	34839	0.9	1.8E+00	P11369	SWISSPROT	RETROVIRUS-RELATED POL POLYPROTEIN [CONTAINS: REVERSE TRANSCRIPTASE; ENDONUCLEASE]
8812	21779	35204	0.46	1.8E+00	P48634	SWISSPROT	LARGE PROLINE-RICH PROTEIN BAT2 (HLA-B-ASSOCIATED TRANSCRIPT 2)
8812	21779	35205	0.46	1.8E+00	P48634	SWISSPROT	LARGE PROLINE-RICH PROTEIN BAT2 (HLA-B-ASSOCIATED TRANSCRIPT 2)
8812	21779	35206	0.46	1.8E+00	P48634	SWISSPROT	LARGE PROLINE-RICH PROTEIN BAT2 (HLA-B-ASSOCIATED TRANSCRIPT 2)
9206	22172	35603	2.21	1.8E+00	O43281	SWISSPROT	EMBRYONAL FYN-ASSOCIATED SUBSTRATE (HEFS)
9530	22493	35941	0.66	1.8E+00	R31042.1	EST_HUMAN	yt72c08.t1 Soares placenta Nb2HP Homo sapiens cDNA clone IMAGE:135278 5'

Page 17 of 546
Table 4
Single Exon Probes Expressed in Bone Marrow

Probe SEQ ID NO:	Exon SEQ ID NO:	ORF SEQ ID NO:	Expression Signal	Most Similar (Top) Hit BLAST E Value	Top Hit Accession No.	Top Hit Database Source	Top Hit Descriptor
9817	22581	36009	0.84	1.8E+00	AW880004.1	EST_HUMAN	QV0-QT0030-070300-148-403 OT0030 Homo sapiens cDNA
10208	23133	36620	0.93	1.8E+00	P27050	SWISSPROT	CHITINASE D PRECURSOR
10846	23568		2.88	1.8E+00	AF111849.1	NT	Homo sapiens PRO0630 mRNA, complete cds
10919	23839		0.76	1.8E+00	P44325	SWISSPROT	CYTIDINE DEAMINASE (CYTIDINE AMINOHYDROLASE) (CDA)
12583	25808		8.17	1.8E+00	AF314254.1	NT	Chlamydomonas reinhardtii alternative oxidase 1 (AOX1) gene, nuclear gene encoding mitochondrial protein
12842	25328		5.63	1.8E+00	9508404	NT	Rattus norvegicus Actin-related protein complex 1b (Apo1b), mRNA
1110	14154	27104	2.43	1.7E+00	Q60114	SWISSPROT	LEVANSUCRASE (BETA-D-FRUCTOFURANOSYL TRANSFERASE) (SUCROSE 6-FRUCTOSYL TRANSFERASE)
2279	15292	28317	3.08	1.7E+00	AL163280.2	NT	Homo sapiens chromosome 21 segment HS21O080
2382	15390	28415	1.87	1.7E+00	AI141087.1	EST_HUMAN	cc43h05.x1 Soares NIH-MIPu_S11 Homo sapiens cDNA clone IMAGE:1678137 3'
4487	17512	30400	0.9	1.7E+00	Q60114	SWISSPROT	LEVANSUCRASE (BETA-D-FRUCTOFURANOSYL TRANSFERASE) (SUCROSE 6-FRUCTOSYL TRANSFERASE)
5694	18789	31959	1.88	1.7E+00	BE083548.1	EST_HUMAN	CM0-BT0282-171289-127-405 BT0282 Homo sapiens cDNA
5694	18789	31960	1.88	1.7E+00	BE083548.1	EST_HUMAN	CM0-BT0282-171289-127-405 BT0282 Homo sapiens cDNA
5855	19040	32238	0.49	1.7E+00	R88748.1	EST_HUMAN	G4846 Fetal heart Homo sapiens cDNA clone G4846 5' end
6133	19210	32436	3.14	1.7E+00	Q9TTR8	SWISSPROT	COUP TRANSCRIPTION FACTOR 1 (COUP-TF1) (COUP-TF1)
6703	19759	33038	0.89	1.7E+00	P35816	SWISSPROT	[PYRUVATE DEHYDROGENASE (LIPONAMIDE)-PHOSPHATASE, CATALYTIC SUBUNIT] (PDPC)
7428	20395	33746	1.01	1.7E+00	Q03703	SWISSPROT	HYPOTHETICAL 38.0 KD PROTEIN IN CAT2-AMD1 INTERGENIC REGION
7428	20395	33747	1.01	1.7E+00	Q03703	SWISSPROT	HYPOTHETICAL 38.0 KD PROTEIN IN CAT2-AMD1 INTERGENIC REGION
8187	21157	34566	0.81	1.7E+00	AF021335.1	NT	Mus musculus T cell receptor gamma locus, TOR gamma 2 and gamma 4 gene clusters
8367	21336	34748	1.21	1.7E+00	8755715	NT	Mus musculus T-cell acute lymphocytic leukemia 1 (Tal1), mRNA
8387	21366	34775	0.54	1.7E+00	BF530630.1	EST_HUMAN	602071917F1 NCI CGAP Brn87 Homo sapiens cDNA clone IMAGE:4214689 5'
8887	21853	35273	0.45	1.7E+00	AF245513.1	NT	Hippoglossus hippoglossus Interferon Inducible Mx protein (Mx) mRNA, complete cds
8974	21940		1.92	1.7E+00	BF308000.1	EST_HUMAN	601894255F1 NIH_MGC_17 Homo sapiens cDNA clone IMAGE:4140084 5'
9052	22018	35442	0.51	1.7E+00	X68063.1	NT	M.musculus Ank-1 mRNA for erythroid ankyrin
9052	22018	35443	0.51	1.7E+00	X68063.1	NT	M.musculus Ank-1 mRNA for erythroid ankyrin
9186	22132	35558	0.43	1.7E+00	U19832.1	NT	Rattus norvegicus SA gene, partial cds
9504	25697	35909	2.37	1.7E+00	O60479	SWISSPROT	HOMEBOX PROTEIN DLX-3
9504	25697	35910	2.37	1.7E+00	O60479	SWISSPROT	HOMEBOX PROTEIN DLX-3
9865	22892		1.4	1.7E+00	AF161380.1	NT	Homo sapiens HSPC382 mRNA, partial cds
10532	23454		0.5	1.7E+00	AW953681.1	EST_HUMAN	EST365751 MAGC resequences, MAGC Homo sapiens cDNA

Page 18 of 546
Table 4
Single Exon Probes Expressed in Bone Marrow

Probe SEQ ID NO:	Exon SEQ ID NO:	ORF SEQ ID NO:	Expression Signal	Most Similar (Top) Hit BLAST E Value	Top Hit Accession No.	Top Hit Database Source	Top Hit Descriptor
12518	26249	31803	1.78	1.7E+00	AI678443.1	EST_HUMAN	tu82d07.x1 NCL_CGAP_Gas4 Homo sapiens cDNA clone IMAGE:2257649 3' similar to contains MSR1.11 MSR1 repetitive element:
2049	15068	28088	14.89	1.6E+00	AF189339.1	NT	Homo sapiens lens epithelium-derived growth factor gene, alternatively spliced, complete cds
2059	15078	28097	3.29	1.6E+00	AF077374.1	NT	Homo sapiens small proline-rich protein (SPRR3) gene, exons 1, 2, and 3 and complete cds
2065	15083	28102	1.62	1.6E+00	Y11344.1	NT	Mus musculus ST6GalNAcII gene, exon 2
2292	15304		1.33	1.6E+00	X88373.1	NT	B. napus gene encoding endo-polygalacturonase
2972	16030	28953	1.88	1.6E+00	W58428.1	EST_HUMAN	zid25f01.1 Soares_fetal heart_NH-1H19W Homo sapiens cDNA clone IMAGE:341688 5' similar to gb:D29805 N-ACETYL LACTOSAMINE SYNTHASE (HUMAN);
4063	17088		7.07	1.6E+00	BF570077.1	EST_HUMAN	602186095T1 NIH_MGC_45 Homo sapiens cDNA clone IMAGE:4310591 3'
4379	17407	30287	1.28	1.6E+00	AF155827.1	NT	Homo sapiens proliferation-associated SNF2-like protein (SMARCA8) mRNA, complete cds
4379	17407	30288	1.28	1.6E+00	AF155827.1	NT	Homo sapiens proliferation-associated SNF2-like protein (SMARCA8) mRNA, complete cds
5120	18130	31006	2.88	1.6E+00	Y11344.1	NT	Mus musculus ST6GalNAcII gene, exon 2
5120	18130	31007	2.88	1.6E+00	Y11344.1	NT	Mus musculus ST6GalNAcII gene, exon 2
5826	19012	32204	2.21	1.6E+00	L04808.1	NT	Brachydanio rerio MHC class II DA-beta-2*01 gene, 3' end
6020	19103	32305	0.82	1.6E+00	AF005631.1	NT	Homo sapiens transglutaminase type I (Tgase1) gene, promoter region
6814	19872	32850	0.83	1.6E+00	BF380703.1	EST_HUMAN	IL2-UT0073-060900-145-E02 UT0073 Homo sapiens cDNA
6868	19921	33217	0.97	1.6E+00	AW204881.1	EST_HUMAN	U1-HB12-eth-b-04-Q-UJ.s1 NCL_CGAP_Sub4 Homo sapiens cDNA clone IMAGE:2727511 3'
7458	20422	33777	2.47	1.6E+00	BE687287.1	EST_HUMAN	RCO-GT0415-200700-032-c10 GT0415 Homo sapiens cDNA
8384	21333		1.1	1.6E+00	Q48378	SWISSPROT	VIRULENCE FACTOR MVIN HOMOLOG
8722	21690	35117	3.42	1.6E+00	AJ297131.1	NT	Mus musculus SIL_MAP_17, CYP_a, SCL & CYP_b genes
9251	22217	35847	0.9	1.6E+00	11437222	NT	Homo sapiens hypothetical protein PRO0871 (PRO0871), mRNA
9251	22217	35848	0.9	1.6E+00	11437222	NT	Homo sapiens hypothetical protein PRO0871 (PRO0871), mRNA
9815	25695	34540	1.54	1.6E+00	X52048.1	NT	M.musculus COL3A1 gene for collagen alpha-1
9815	25695	34541	1.54	1.6E+00	X52048.1	NT	M.musculus COL3A1 gene for collagen alpha-1
9845	25872		0.48	1.6E+00	AF043468.1	NT	Thermotoga maritima ethanolicus D-xylose-binding protein (xylF) gene, complete cds
10091	23017	38493	1.23	1.6E+00	T41290.1	EST_HUMAN	ph6b6_19/1TV Outward Aki-primed hncDNA library/Homo sapiens cDNA clone ph6b6_19/1TV
10506	23428	36925	0.48	1.6E+00	AF121361.1	NT	Drosophila melanogaster signal transducing adaptor protein (STAM), serine threonine kinase lal (LAL), and zinc finger protein (DNZ1) genes, complete cds
10545	23467	36861	1.12	1.6E+00	AW835844.1	EST_HUMAN	QV4-LT0016-080200-100-007 LT0016 Homo sapiens cDNA
10545	23467	36862	1.12	1.6E+00	AW835844.1	EST_HUMAN	QV4-LT0016-080200-100-007 LT0016 Homo sapiens cDNA
10709	23631	37126	0.78	1.6E+00	AF037352.1	NT	Mus musculus T cell receptor gamma locus, TCR gamma 1 and gamma 3 gene clusters
11123	24083	37609	1.73	1.6E+00	P54817	SWISSPROT	CAPSID PROTEIN P40 (CONTAINS: ASSEMBLIN (PROTEASE); CAPSID ASSEMBLY PROTEIN)
11192	19103	32305	5.78	1.6E+00	AF005631.1	NT	Homo sapiens transglutaminase type I (Tgase1) gene, promoter region
12017	24894	38491	2.83	1.6E+00	AF104313.1	NT	Homo sapiens unknown mRNA

Page 19 of 546
Table 4

Single Exon Probes Expressed in Bone Marrow

Probe SEQ ID NO:	Exon SEQ ID NO:	ORF SEQ ID NO:	Expression Signal	Most Similar (Top) Hit BLAST E Value	Top Hit Accession No.	Top Hit Database Source	Top Hit Descriptor
35	13155	26056	3.54	1.5E+00	U53449.1	NT	Rattus norvegicus Jun dimerization protein 2 (jdp-2) mRNA, complete cds
234	13334	26258	2.63	1.5E+00	AE002201.2	NT	Chlamydomonas reinhardtii AR39, section 32 of 94 of the complete genome
622	13687		2.04	1.5E+00	6752861	NT	Mus musculus a disintegrin and metalloproteinase domain (ADAM) 15 (metabinklin) (Adam15), mRNA
2420	15427	28450	3.2	1.5E+00	AJ131402.1	NT	Potato virus A RNA complete genome, isolate U
2522	15525	28547	2.17	1.5E+00	6878350	NT	Mus musculus T-cell lymphoma invasion and metastasis 1 (Tiam1), mRNA
3155	15427	28450	2.96	1.5E+00	AJ131402.1	NT	Potato virus A RNA complete genome, isolate U
3386	16435	29382	0.92	1.5E+00	AE001945.1	NT	Dendrococcus radiodurans R1 section 82 of 229 of the complete chromosome 1
5817	18907	32090	0.82	1.5E+00	AI655301.1	EST_HUMAN	h12f10.x1 NCI_CGAP_G08 Homo sapiens cDNA clone IMAGE:2240587 3' similar to TR:O00237 O00237 HKF-1.;
5817	18907	32091	0.82	1.5E+00	AI655301.1	EST_HUMAN	h12f10.x1 NCI_CGAP_G08 Homo sapiens cDNA clone IMAGE:2240587 3' similar to TR:O00237 O00237 HKF-1.;
6546	19607	32899	2.71	1.5E+00	R17878.1	EST_HUMAN	yg10e02.r1 Soares infant brain 1NIB Homo sapiens cDNA clone IMAGE:31693 5'
7335	20308		1.48	1.5E+00	BE783568.1	EST_HUMAN	601478745F1 NIH_MGC_68 Homo sapiens cDNA clone IMAGE:3881555 5'
7368	20338	33688	16.72	1.5E+00	P47179	SWISSPROT	HYPOTHETICAL 118.4 KD PROTEIN IN BAT2-DAL5 INTERGENIC REGION PRECURSOR
7368	20338	33689	16.72	1.5E+00	P47179	SWISSPROT	HYPOTHETICAL 118.4 KD PROTEIN IN BAT2-DAL5 INTERGENIC REGION PRECURSOR
7568	20531	33890	0.6	1.5E+00	AA889259.1	EST_HUMAN	ak28f10.s1 Soares testis_NHT Homo sapiens cDNA clone IMAGE:1407115 3'
7850	20798	34172	0.73	1.5E+00	AI003254.1	EST_HUMAN	an07b11.s1 Strategene schizos brain S11 Homo sapiens cDNA clone IMAGE:1684893 3' similar to gb:595936 SEROTRANSFERRIN PRECURSOR (HUMAN);
8144	21081		0.67	1.5E+00	AB039887.1	NT	Homo sapiens WDR4 gene for WD repeat protein, complete cds
8469	21428	34845	0.91	1.5E+00	BE887446.1	EST_HUMAN	601509566F1 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:3911181 5'
8514	21482	34896	0.48	1.5E+00	AB040887.1	NT	Homo sapiens mRNA for KIAA1454 protein, partial cds
8691	21957	35382	1.08	1.5E+00	K02138.1	NT	Mouse germline IgM chain gene, mu-delta region
8870	22335		0.47	1.5E+00	AB039516.1	NT	Homo sapiens hGPIb alpha gene for platelet glycoprotein Ib alpha, complete cds
9488	22453	36893	0.55	1.5E+00	BF217818.1	EST_HUMAN	yj03h01.r1 Soares placenta Nb2HP Homo sapiens cDNA clone IMAGE:4095135 5'
9838	22774	36229	0.91	1.5E+00	R61928.1	EST_HUMAN	QV3-CT0192-201099-008-009 CT0182 Homo sapiens cDNA
9893	22920	36387	1.09	1.5E+00	AW375897.1	EST_HUMAN	RC0-TN0078-150900-034-g05 TN0078 Homo sapiens cDNA
10219	23144	36833	6.14	1.5E+00	BF376754.1	EST_HUMAN	602035771F1 NCI_CGAP_Bim64 Homo sapiens cDNA clone IMAGE:4183865 5'
10412	23334		1.42	1.5E+00	BF337944.1	EST_HUMAN	z38g06.r1 Soares retina N2b4HR Homo sapiens cDNA clone IMAGE:361308 5'
10556	23478	36972	2.92	1.5E+00	AA017689.1	EST_HUMAN	z38g06.r1 Soares retina N2b4HR Homo sapiens cDNA clone IMAGE:361308 5'
10556	23478	36973	2.92	1.5E+00	AA017689.1	EST_HUMAN	z38g06.r1 Soares retina N2b4HR Homo sapiens cDNA clone IMAGE:361308 5'
11727	24913	39190	5.98	1.5E+00	AL134197.1	EST_HUMAN	DKFZp647P243.s1 547 (synonym: hibr1) Homo sapiens cDNA clone DKFZp647P243 3'
11869	24751		6.39	1.5E+00	X07390.1	NT	Malva mitochondrial tRNA-Ser gene and tRNA-Phe pseudogene
12112	24982	38583	1.53	1.5E+00	BE257552.1	EST_HUMAN	601109621F1 NIH_MGC_16 Homo sapiens cDNA clone IMAGE:3350477 5'

Table 4

Single Exon Probes Expressed in Bone Marrow

Probe SEQ ID NO:	Exon SEQ ID NO:	ORF SEQ ID NO:	Expression Signal	Most Similar (Top) BLAST E Value	Top Hit Accession No.	Top Hit Database Source	Top Hit Descriptor
12862	25151		1.87	1.5E+00	8753287	NT	Mus musculus caspase 8 associated protein 2 (Casp8ap2), mRNA
12510	25899	31420	1.51	1.5E+00	D63480.1	NT	Human mRNA for KIAA0148 gene, partial cds
12723	25376		4.94	1.5E+00	AL445085.1	NT	Thermoplasma acidophilum complete genome, segment 3/5
12821	25443		1.33	1.5E+00	8878482	NT	Rattus norvegicus 5 - Lipoxigenase (ALOX5), mRNA
32	13152	26052	1.41	1.4E+00	7661885	NT	Homo sapiens DKFZP588M0122 protein (DKFZP588M0122), mRNA
32	13152	26053	1.41	1.4E+00	7661885	NT	Homo sapiens DKFZP588M0122 protein (DKFZP588M0122), mRNA
2285	15288		0.9	1.4E+00	AF053357.1	NT	Helicobacter pylori glutamine synthetase (ghsA) gene, complete cds
2341	15351		8.77	1.4E+00	U67822.1	NT	Ovis aries prion protein gene, complete cds
2677	15673	26863	1.83	1.4E+00	X74463.1	NT	Human papillomavirus type 7 genomic DNA
2784	15776	28784	2.76	1.4E+00	AF064564.2	NT	Fugu rubripes neurofibromatosis type 1 (NF1), A-kinase anchor protein (AKAP84), BAW protein (BAW), and WSB1 protein (WSB1) genes, complete cds
2784	15776	28785	2.75	1.4E+00	AF064564.2	NT	Fugu rubripes neurofibromatosis type 1 (NF1), A-kinase anchor protein (AKAP84), BAW protein (BAW), and WSB1 protein (WSB1) genes, complete cds
3345	16398		0.63	1.4E+00	5433733	NT	Homo sapiens Mad4 homolog (MAD4) mRNA
4284	17313	30181	1.38	1.4E+00	AW900455.1	EST_HUMAN	CMO-NN1005-140300-288-H08 NN1005 Homo sapiens cDNA
4284	17313	30182	1.38	1.4E+00	AW900455.1	EST_HUMAN	CMO-NN1005-140300-288-H08 NN1005 Homo sapiens cDNA
4580	16144	29056	0.82	1.4E+00	AE002324.2	NT	Chlamydia muridarum, section 55 of 85 of the complete genome
4580	16144	29057	0.82	1.4E+00	AE002324.2	NT	Chlamydia muridarum, section 55 of 85 of the complete genome
4815	17636		1.71	1.4E+00	BF081547.1	EST_HUMAN	602156687F1 NIH_MGC_83 Homo sapiens cDNA clone IMAGE:4297556 5'
6446	18648	31461	1.58	1.4E+00	AW054878.1	EST_HUMAN	w45g07 x1 NCI_CGAP_Pan1 Homo sapiens cDNA clone IMAGE:2510460 3'
5808	18702		5.52	1.4E+00	AB032983.1	NT	Homo sapiens mRNA for KIAA1157 protein, partial cds
8410	19478	32725	3.06	1.4E+00	Q13472	SWISSPROT	DNA TOPOISOMERASE III ALPHA
8427	25998		4.2	1.4E+00	AB020712.1	NT	Homo sapiens mRNA for KIAA0905 protein, complete cds
8552	19613	32875	2.65	1.4E+00	Q92777	SWISSPROT	SYNAPSIN II
8552	19613	32876	2.65	1.4E+00	Q92777	SWISSPROT	SYNAPSIN II
6781	19638	33119	0.56	1.4E+00	BE007870.1	EST_HUMAN	QVO-BN0148-050500-215-b11 BN0148 Homo sapiens cDNA
6781	19638	33120	0.56	1.4E+00	BE007870.1	EST_HUMAN	QVO-BN0148-050500-215-b11 BN0148 Homo sapiens cDNA
6886	20122	33436	0.86	1.4E+00	AW893057.1	EST_HUMAN	CM3-NN0008-300300-132-b12 NN0008 Homo sapiens cDNA
							Homo sapiens cavedin-1/-2 locus, Config1, D7S522, genes CAV2 (exons 1, 2a, and 2b), CAV1 (exons 1 and 2)
7503	20488	33829	2.02	1.4E+00	AJ133289.1	NT	ha23f05.x1 NCI_CGAP_CML1 Homo sapiens cDNA clone IMAGE:2318873 3' similar to contains Alu repetitive element
7521	20486	33848	1.15	1.4E+00	AW487780.1	EST_HUMAN	LAMININ BETA-2 CHAIN PRECURSOR (S-LAMININ)
7582	20544	33904	0.58	1.4E+00	P55288	SWISSPROT	LAMININ BETA-2 CHAIN PRECURSOR (S-LAMININ)
7582	20544	33905	0.58	1.4E+00	P55288	SWISSPROT	LAMININ BETA-2 CHAIN PRECURSOR (S-LAMININ)

Page 21 of 546
Table 4

Single Exon Probes Expressed in Bone Marrow

Probe SEQ ID NO:	Exon SEQ ID NO:	ORF SEQ ID NO:	Expression Signal	Most Similar (Top) Hit BLAST E Value	Top Hit Accession No.	Top Hit Database Source	Top Hit Descriptbr
7611	20571	33935	0.65	1.4E+00	Q80905	SWISSPROT	MINOR CAPSID PROTEIN L2
8678	21646			1.4E+00	P07983	SWISSPROT	GLUCOAMYLASE PRECURSOR (GLUCAN 1,4-ALPHA-GLUCOSIDASE) (1,4-ALPHA-D-GLUCAN GLUCOHYDROLASE)
9146	22112		4.6	1.4E+00	AJ271735.1	NT	Homo sapiens Xq pseudautosomal region; segment 1/2
9449	22413	35849	1.56	1.4E+00	R20459.1	EST_HUMAN	Y63312.1 Soares Infant brain 1N1B Homo sapiens cDNA clone IMAGE:34945 5'
9552	22514	35865	3.46	1.4E+00	BE064687.1	EST_HUMAN	RC1-BT0313-301289-012-605 BT0313 Homo sapiens cDNA
9587	22549	36000	0.45	1.4E+00	AF134844.1	NT	Sceloporus undulatus ornithine transcarbamylase (OTC) mRNA, complete cds
10568	23490	36982	0.89	1.4E+00	BF575545.1	EST_HUMAN	602133135F1 NIH_MGC_81 Homo sapiens cDNA clone IMAGE:4288137 5'
10613	23535	37032	0.69	1.4E+00	BE146374.1	EST_HUMAN	IL5-HT0198-291088-008-C04 HT0198 Homo sapiens cDNA
10613	23535	37033	0.69	1.4E+00	BE146374.1	EST_HUMAN	IL5-HT0198-291088-008-C04 HT0198 Homo sapiens cDNA
10892	23812	37319	0.92	1.4E+00	D63441.1	NT	Pandorina codemianae chloroplast rbcL gene for ribulose biphosphatase, partial cds
10892	23812	37320	0.92	1.4E+00	D63441.1	NT	Pandorina codemianae chloroplast rbcL gene for ribulose biphosphatase, partial cds
11559	24499	38055	4.43	1.4E+00	AB006882.1	NT	Homo sapiens APECD mRNA for AIRE-1, complete cds
11733	24619	38106	2.83	1.4E+00	BE962107.2	EST_HUMAN	601655194R1 NIH_MGC_65 Homo sapiens cDNA clone IMAGE:3845805 3'
11733	24619	38197	2.83	1.4E+00	BE962107.2	EST_HUMAN	601655194R1 NIH_MGC_65 Homo sapiens cDNA clone IMAGE:3845805 3'
11756	24684	38263	2.15	1.4E+00	U30790.1	NT	Pneumocystis carinii f. sp. ratii guanine nucleotide binding protein alpha subunit (pog1) gene, complete cds
11756	24684	38264	2.15	1.4E+00	U30790.1	NT	Pneumocystis carinii f. sp. ratii guanine nucleotide binding protein alpha subunit (pog1) gene, complete cds
12360	25827		1.7	1.4E+00	AL161500.2	NT	Arabidopsis thaliana DNA chromosome 4, contig fragment No. 12
13098	25602		1.4	1.4E+00	7657824	NT	Homo sapiens staufen (Drosophila, RNA-binding protein) homolog 2 (STAU2), mRNA
572	13641		1.56	1.3E+00	Z73640.1	NT	M.mucoso gene encoding 4-Dihydroxyethyl-lysine dehydrogenase
903	13958	26915	3.12	1.3E+00	AJ271192.1	NT	Cantharellus sp. partial 26S rRNA gene, isolate Tibet
1131	14174		32.4	1.3E+00	Y18213.1	NT	Homo sapiens putative psittacine pseudogene for hair keratin, exons 2 to 7
1300	14336	27299	19.22	1.3E+00	4507998	NT	Homo sapiens zinc finger protein 157 (ZNF157) mRNA
1300	14336	27300	19.22	1.3E+00	4507998	NT	Homo sapiens zinc finger protein 157 (ZNF157) mRNA
1358	14393		0.99	1.3E+00	U61730.2	NT	Cox laeyriae-101 dihydrodipicolinate synthase (daps) gene, complete cds
1614	14946		2.69	1.3E+00	AE002338.2	NT	Chlamydia muridarum, section 86 of 85 of the complete genome
							Cyprinus carpio MRPs and MASPb genes for mannose-binding lectin-associated serine protease (MASP)
2255	15289		1.22	1.3E+00	AB030447.1	NT	and MASP-related protein, complete cds
2416	15422	28445	1.03	1.3E+00	P25391	SWISSPROT	LAMININ ALPHA-1 CHAIN PRECURSOR (LAMININ A CHAIN)
2557	15559		2.6	1.3E+00	BE966755.2	EST_HUMAN	601661233R1 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:3915945 3'
2950	16007	29631	0.67	1.3E+00	6755621	NT	Mus musculus alpha-spectrin 1, erythroid (Sptn1), mRNA

Page 22 of 546
Table 4[illegible]

Page 23 of 546
Table 4

Single Exon Probes Expressed in Bone Marrow

Probe SEQ ID NO:	Exon SEQ ID NO:	ORF SEQ ID NO:	Expression Signal	Most Similar (Top) Hit BLAST E Value	Top Hit Accession No.	Top Hit Database Source	Top Hit Descriptor
10286	23211		0.64	1.3E+00	AI569844.1	EST_HUMAN	igT7a12x1 NCI CGAP_UH1 Homo sapiens cDNA clone IMAGE:2214814 3' similar to gbX14723
10508	23431	36927	0.45	1.3E+00	AF061251.1	NT	CLUSTERIN PRECURSOR (HUMAN);
10509	23431	36928	0.45	1.3E+00	AF061251.1	NT	Escherichia coli serotype O157:H7 O antigen gene cluster
10574	23486	36989	1.24	1.3E+00	AE004382.1	NT	Escherichia coli serotype O157:H7 O antigen gene cluster
10591	23513	37008	1.53	1.3E+00	M28953.1	NT	Vibrio cholerae chromosome II, section 49 of 93 of the complete chromosome
10948	23868		0.85	1.3E+00	AL163302.2	NT	Campylobacter jejuni kanamycin phosphotransferase (aphA-7) gene, complete cds
10975	23895	37409	0.49	1.3E+00	AI890846.1	EST_HUMAN	Homo sapiens chromosome 21 segment HS21C102
10987	23907		0.43	1.3E+00	8923637	NT	hs32e10.x1 NCI CGAP_G08 Homo sapiens cDNA clone IMAGE:2498822 3' similar to SW:TRXB_HUMAN
11052	24015		3.63	1.3E+00	Q14117	SWISSPROT	Q16881 THIOREDOXIN REDUCTASE;
11252	24205	37727	2.35	1.3E+00	P25299	SWISSPROT	Homo sapiens hypodermal protein FLJ20707 (FLJ20707), mRNA
11273	24225	37751	1.71	1.3E+00	Z18892.2	NT	DIHYDROPYRIMIDINASE (DHPASE) (HYDANTOINASE) (DHP)
11867	24749	38331	2.61	1.3E+00	D42042.1	NT	MRNA 3'-END PROCESSING PROTEIN RNA15
11945	24825	38420	2.1	1.3E+00	Z98982.1	NT	Mus musculus desmin gene
12001	24878		1.55	1.3E+00	L31891.1	NT	Human mRNA for KIAA0085 gene, partial cds
12498	25239		3.37	1.3E+00	AF187873.1	NT	Bacillus subtilis genomic DNA 23.8kb fragment
12669	25344	31762	6.47	1.3E+00	BF348043.1	EST_HUMAN	Arabidopsis thaliana 3-ketacyl-acyl carrier protein synthase III (KAS III) mRNA, complete cds
12679	25724		3.24	1.3E+00	P33464	SWISSPROT	Arabidopsis thaliana 3-ketacyl-acyl carrier protein synthase III (KAS III) mRNA, complete cds
12765	25407		1.68	1.3E+00	AF187035.1	NT	Cavia porcellus inwardly-rectifying potassium channel Kir2.2 (KCNU12) gene, complete cds
651	13717	26639	13.29	1.2E+00	AA678246.1	EST_HUMAN	602023185F1 NCI CGAP_Bm67 Homo sapiens cDNA clone IMAGE:4158452 5'
824	13882	26833	1.37	1.2E+00	P05228	SWISSPROT	E1 GLYCOPROTEIN PRECURSOR (MATRIX GLYCOPROTEIN) (MEMBRANE GLYCOPROTEIN)
824	13882	26834	1.37	1.2E+00	P05228	SWISSPROT	Stumra illum cytochrome b gene, complete cds; mitochondrial gene for mitochondrial product
824	13882	26835	1.37	1.2E+00	P05228	SWISSPROT	z12x08.s1 Soares_fetal_liver_spleen_INFLS_S1 Homo sapiens cDNA clone IMAGE:431535 3'
879	13934		0.95	1.2E+00	8924234	NT	HISTIDINE-RICH PROTEIN PRECURSOR (CLONE PFHRP-III)
1166	14206	27160	7.4	1.2E+00	AF080245.2	NT	HISTIDINE-RICH PROTEIN PRECURSOR (CLONE PFHRP-II)
1210	14248	27206	1.43	1.2E+00	AJ252242.1	NT	HISTIDINE-RICH PROTEIN PRECURSOR (CLONE PFHRP-II)
1210	14248	27207	1.43	1.2E+00	AJ252242.1	NT	HISTIDINE-RICH PROTEIN PRECURSOR (CLONE PFHRP-II)
2025	15045	28058	1.06	1.2E+00	AF140631.1	NT	Homo sapiens hypodermal protein PRO3077 (PRO3077), mRNA
3127	16184	28093	0.99	1.2E+00	AB020681.1	NT	Elasid clefata sesquiterpene synthase mRNA, complete cds
3179	16234	28151	6.14	1.2E+00	AL161563.2	NT	pea seed-borne mosaic virus complete genome
3179	16234	28152	6.14	1.2E+00	AL161563.2	NT	pea seed-borne mosaic virus complete genome
3299	16352		3.07	1.2E+00	P64910	SWISSPROT	Homo sapiens G-protein coupled receptor 14 (GPR14) gene, complete cds
3337	16417	28343	0.78	1.2E+00	AF188740.1	NT	Homo sapiens mRNA for KIAA0874 protein, partial cds

Page 24 of 546
Table 4

Single Exon Probes Expressed in Bone Marrow

Probe SEQ ID NO:	Exon SEQ ID NO:	ORF SEQ ID NO:	Expression Signal	Most Similar (Top) Hit BLAST E Value	Top Hit Accession No.	Top Hit Database Source	Top Hit Descriptor
3725	16767	29678	8.46	1.2E+00	U75902.1	NT	Mus musculus subtilisin-like serine protease LPC (PC7) gene, exons 1 to 9, partial cds
4012	17051	29657	1.78	1.2E+00	BF373570.1	EST_HUMAN	MRO-FT0175-050900-203-g08_1 FT0175 Homo sapiens cDNA
4327	18417	28343	1.09	1.2E+00	AF188740.1	NT	Homo sapiens LH33 gene, intron 2
4498	17823		2.12	1.2E+00	M87060.1	NT	Rattus rattus cardiac AE3 gene, exons 1-23
4551	17574	30464	1.33	1.2E+00	AL161509.2	NT	Arabidopsis thaliana DNA chromosome 4, contig fragment No. 21
4592	17813	30507	2.1	1.2E+00	AF169466.1	NT	Homo sapiens post-synaptic density 95 (DLG4) gene, complete cds
4619	17640		6.6	1.2E+00	Y08230.1	NT	T. phaeatum chloroplast rbcL gene, partial
5052	18064	30942	2.36	1.2E+00	P05228	SWISSPROT	HISTIDINE-RICH PROTEIN PRECURSOR (GLONE PHRP-III)
5052	18064	30943	2.36	1.2E+00	P05228	SWISSPROT	HISTIDINE-RICH PROTEIN PRECURSOR (GLONE PHRP-III)
5052	18064	30944	2.36	1.2E+00	P05228	SWISSPROT	HISTIDINE-RICH PROTEIN PRECURSOR (GLONE PHRP-III)
5512	18812	31544	1.08	1.2E+00	U20760.1	NT	Human extracellular calcium-sensing receptor mRNA, complete cds
5634	18730	31892	1.89	1.2E+00	AW813276.1	EST_HUMAN	MR3-ST0191-140200-013-c05 ST0191 Homo sapiens cDNA
5894	18982		0.57	1.2E+00	X61878.1	NT	Calicivirus cDNA for orf1, orf2 and orf3
5974	19059	32280	0.78	1.2E+00	AF016052.1	NT	Homo sapiens zinc finger protein ZNF191 (ZNF191) gene, complete cds
6276	19248	32580	2.39	1.2E+00	X74895.1	NT	D. hydei av1 repeat cluster DNA, fragment D
6398	19407	32648	4.12	1.2E+00	BE003113.1	EST_HUMAN	QV4-BN0080-270400-190-a03 BN0080 Homo sapiens cDNA
6423	19490	32741	1.41	1.2E+00	X89084.1	NT	G-glutamyl pta gene and ackA gene
6423	19490	32742	1.41	1.2E+00	X89084.1	NT	G-glutamyl pta gene and ackA gene
6467	19532	32780	38.6	1.2E+00	AA769254.1	EST_HUMAN	ah84g12.s1 Soares testis_NHT Homo sapiens cDNA clone 1322374 3'
6575	19635	32801	0.8	1.2E+00	N33295.1	EST_HUMAN	y93b12.s1 Soares melanocyte 2NbtHM Homo sapiens cDNA clone IMAGE:273589 3' similar to gb MB7835 HUMAAU1472 Human carcinoma cell-derived Alu RNA transcript4 (rRNA); gb J04970
6650	19708	32884	0.71	1.2E+00	P17671	SWISSPROT	CARBOXYPEPTIDASE M PRECURSOR (HUMAN);
6654	19711	32888	2.01	1.2E+00	AW813276.1	EST_HUMAN	ECDYSONE-INDUCIBLE PROTEIN E75-A
7100	20034	33337	1.17	1.2E+00	AB029010.1	NT	MR3-ST0191-140200-013-c05 ST0191 Homo sapiens cDNA
7113	20047	33349	2.68	1.2E+00	AJ002141.1	NT	Homo sapiens mRNA for KIAA1087 protein, partial cds
7465	20431		0.64	1.2E+00	AJ271735.1	NT	Mus musculus DSPP gene
7610	20685	33934	1.64	1.2E+00	AV734585.1	EST_HUMAN	Homo sapiens Xq pseudobacterial region; segment 1/2
7912	20855	34243	2.6	1.2E+00	X74207.1	NT	AV734585 cda Homo sapiens cDNA clone cdaAAAF-103 5'
8122	21059	34457	0.53	1.2E+00	BE787946.1	EST_HUMAN	Lactis pyD and pyF genes
8915	21881	35307	3.24	1.2E+00	AB030300.1	NT	601481761F1 NIH_MGC_68 Homo sapiens cDNA clone IMAGE:3884270 5'
							Homo sapiens mRNA for KIAA1204 protein, partial cds
							ALPHA,ALPHA-TREHALOSE-PHOSPHATE SYNTHASE (UDP-FORMING) 123 KD SUBUNIT (TREHALOSE-6-PHOSPHATE SYNTHASE)(UDP-GLUCOSE-GLUCOSEPHOSPHATE
9010	21976	35398	0.65	1.2E+00	P39427	SWISSPROT	GLUCOSYL TRANSFERASE)

Page 25 of 546
Table 4

Single Exon Probes Expressed in Bone Marrow

Probe SEQ ID NO:	Exon SEQ ID NO:	ORF SEQ ID NO:	Expression Signal	Most Similar (Top) Hit BLAST E Value	Top Hit Accession No.	Top Hit Database Source	Top Hit Descriptor
9228	22194		0.58	1.2E+00	7708271	NT	Homo sapiens CGI-30 protein (LOC51611), mRNA
9380	22345	35777	1.88	1.2E+00	AW977210.1	EST_HUMAN	MR2-CT0222-201089-001-e07 CT0222 Homo sapiens cDNA
9504	22556	36008	0.47	1.2E+00	H48599.1	EST_HUMAN	y80a06.r1 Soares fetal liver spleen 1NF1.S Homo sapiens cDNA clone IMAGE:202066 5'
9753	22694	36150	3.63	1.2E+00	Z32850.1	NT	R.communis gene for pyrophosphate-dependent phosphofructokinase beta subunit
9884	22891	36953	1.71	1.2E+00	D11745.1	EST_HUMAN	HUMHM01A01 Liver HepG2 cell line. Homo sapiens cDNA clone hnm01a01
10291	23216	36700	3.62	1.2E+00	X58832.1	NT	H. sapiens ENO3 gene for muscle specific enolase
10689	23611		0.73	1.2E+00	AB006666.1	NT	Homo sapiens Ikbho gene, exon 1
11674	24840	38219	1.88	1.2E+00	AW917817.1	EST_HUMAN	PMO-ST0284-161189-001-d01 ST0284 Homo sapiens cDNA
11713	24676		7.08	1.2E+00	BE180761.1	EST_HUMAN	PM1-HT0422-160200-007-g10 HT0422 Homo sapiens cDNA
11785	23040	37482	2.59	1.2E+00	U50147.1	NT	Rattus norvegicus synapse-associated protein 102 mRNA, complete cds
12466	25799	31524	18.09	1.2E+00	AL163203.2	NT	Homo sapiens chromosome 21 segment HS21C0003
12487	25232		2.7	1.2E+00	AP001615.1	NT	Bacillus halodurans genomic DNA, section 9/14
464	13537	28463	1.43	1.1E+00	D86880.1	NT	Human mRNA for KIAA0227 gene, partial cds
1777	14808	27792	1.97	1.1E+00	AW965933.1	EST_HUMAN	QV0-BN0042-170300-163-g12 BN0042 Homo sapiens cDNA
1915	14939	27935	1.08	1.1E+00	AW575889.1	EST_HUMAN	UJHF-BR0P-ek-4-02-0-UI.s1 NIH_MGC_52 Homo sapiens cDNA clone IMAGE:3074834 3'
3341	16392	28313	7.34	1.1E+00	AL163213.2	NT	Homo sapiens chromosome 21 segment HS21C0013
3341	16392	28314	7.34	1.1E+00	AL163213.2	NT	Homo sapiens chromosome 21 segment HS21C0013
3498	16545	28471	0.79	1.1E+00	8922841	NT	Homo sapiens hypothetical protein FLJ10749 (FLJ10749), mRNA
3583	16638	28558	3.34	1.1E+00	AI808380.1	EST_HUMAN	wf54h11.x1 Soares NFL_T_GBC_S1 Homo sapiens cDNA clone IMAGE:2359481 3' similar to SW-F531_HUMAN Q12888 P63-BINDING PROTEIN 63BP1 ;
3731	16773	29684	1.52	1.1E+00	AE003886.1	NT	Xylella fastidiosa, section 32 of 229 of the complete genome
3731	16773	29685	1.52	1.1E+00	AE003886.1	NT	Xylella fastidiosa, section 32 of 229 of the complete genome
3831	16871		0.86	1.1E+00	X85374.1	NT	H. parahaemolyticus hphIM(A), hphIM(C), hphIR and menB genes
3964	17004	28918	0.87	1.1E+00	8622841	NT	Homo sapiens hypothetical protein FLJ10749 (FLJ10749), mRNA
4050	17087	28983	1.02	1.1E+00	6755205	NT	Mus musculus proteasome (prosome, macropain) subunit, beta type 7 (Pamb7), mRNA
4243	17272		6.78	1.1E+00	5935331	NT	R. uniconis complete mitochondrial genome
5028	18042	30925	3.15	1.1E+00	U18466.1	NT	African swine fever virus, complete genome
5085	18095	30971	3.91	1.1E+00	U94740.1	NT	Emeticella nidulans sterigmatocystin biosynthetic gene cluster: (stcA), (stcB), (stcC), (stcE), (stcF), (stcI), (stcJ), (stcK), (stcL), (stcO), (stcP), (stcQ), (stcS), (stcT), (stcU), (stcV) and (stcW) genes, complete cds
5109	18119	30983	1.04	1.1E+00	X78425.1	NT	E. faecalis pbp5 gene
5289	18276	31139	0.9	1.1E+00	6680080	NT	Homo sapiens putative GR6 protein (GR6), mRNA
5380	18484	31359	1.53	1.1E+00	8978530	NT	Rattus norvegicus Aquaporin 4 (Aqp4), mRNA
5395	18730	31861	13.39	1.1E+00	BE560184.1	EST_HUMAN	601652776R1 NIH_MGC_58 Homo sapiens cDNA clone IMAGE:3825835 3'

Page 26 of 546
Table 4

Single Exon Probes Expressed in Bone Marrow

Probe SEQ ID NO:	Exon SEQ ID NO:	ORF SEQ ID NO:	Expression Signal	Most Similar (Top) Hit BLAST E Value	Top Hit Accession No.	Top Hit Database Source	Top Hit Descriptor
6715	18009	31987	0.99	1.1E+00	A138582.1	EST_HUMAN	q085c03.x1 Soares_testis_NHT Homo sapiens cDNA clone IMAGE:1736260 3'
6211	19285	32517	1.83	1.1E+00	11419739	NT	Homo sapiens solute carrier family 6 (neurotransmitter transporter), member 14 (SLC8A14), mRNA
6405	19473	32721	0.94	1.1E+00	AF197961.1	NT	Macgregoria pulchra cytochrome b gene, complete cds; mitochondrial gene for mitochondrial product
6547	18008	32870	0.83	1.1E+00	R00037.1	EST_HUMAN	ye80c03.r1 Soares fetal liver spleen TNF α Homo sapiens cDNA clone IMAGE:124924 5'
6876	19029	33228	1.21	1.1E+00	AJ404004.1	NT	Mus musculus mRNA for ER protein 58 (EP58 gene)
7468	20434		0.68	1.1E+00	AF101091.1	NT	Homo sapiens collagen type XI alpha-1 (COL11A1) gene, exons 25 through 28
7514	20479	33840	0.82	1.1E+00	X55981.1	NT	Maize mRNA for endase (2-phospho-D-glycerate hydrolase)
7732	20687	34050	1.83	1.1E+00	Z72338.1	NT	Herpes simplex virus type 1 (strain KOS) UL41 gene
7732	20687	34051	1.83	1.1E+00	Z72338.1	NT	Herpes simplex virus type 1 (strain KOS) UL41 gene
7754	20707	34076	8.13	1.1E+00	AL161588.2	NT	Arabidopsis thaliana DNA chromosome 4, contig fragment No. 84
7833	25691	34158	0.94	1.1E+00	11067960	NT	Mus musculus silent mating type information regulation 2, (S.cerevisiae, homolog)-like (Sir2), mRNA
8470	21439	34857	2.85	1.1E+00	BF693996.1	EST_HUMAN	602082582F1 NIH_MGC_81 Homo sapiens cDNA clone IMAGE:4246628 5'
8563	21531	34961	0.73	1.1E+00	A1478339.1	EST_HUMAN	hm39h11.x1 NCI_CGAP_K3d11 Homo sapiens cDNA clone IMAGE:2160549 3'
9089	22055	35479	0.63	1.1E+00	AB003088.1	NT	Acetabularia caliculatus mitochondrial COXI-like gene
9167	22133	35559	1.42	1.1E+00	S80750.1	NT	VH=anti-cytomegalovirus glycoprotein B antibody 4D4 heavy chain variable region [human, mRNA Partial, 375 nt]
9278	22244	35673	0.45	1.1E+00	A1079948.1	EST_HUMAN	023403.x1 Soares_NHMPu_S1 Homo sapiens cDNA clone IMAGE:1677249 3'
9792	21115		0.64	1.1E+00	BE384978.1	EST_HUMAN	601276278F1 NIH_MGC_20 Homo sapiens cDNA clone IMAGE:3617418 5'
9886	22913	36378	0.85	1.1E+00	AJ245772.1	NT	Mus musculus mRNA for stretch responsive muscle (X-chromosome) protein (Srmx gene)
10040	22987		0.69	1.1E+00	Y12227.1	NT	Arabidopsis thaliana DNA, 24 kb surrounding PFL locus
10130	23056	36634	0.95	1.1E+00	L76301.1	NT	Yersinia pseudotuberculosis psaeE, psaeF, adhesin (psaA), chaperone (psaB), and usher (psaC) genes, complete cds
10192	23117	36601	1.38	1.1E+00	AB023151.1	NT	Homo sapiens mRNA for KIAA0934 protein, partial cds
10267	23222	36706	6.66	1.1E+00	AL161515.2	NT	Arabidopsis thaliana DNA chromosome 4, contig fragment No. 27
10357	23281	36757	20.52	1.1E+00	6754021	NT	Mus musculus guanine nucleotide binding protein (G protein), gamma 3 subunit (Gng3), mRNA
10867	23787	37287	1.09	1.1E+00	P73769	SWISSPROT	DNA MISMATCH REPAIR PROTEIN MUTS
10889	23909	37422	0.63	1.1E+00	BF343844.1	EST_HUMAN	602014488F1 NCI_CGAP_Bm64 Homo sapiens cDNA clone IMAGE:4150508 5'
10889	23909	37423	0.63	1.1E+00	BF343844.1	EST_HUMAN	602014488F1 NCI_CGAP_Bm64 Homo sapiens cDNA clone IMAGE:4150508 5'
11012	23977	37502	2.03	1.1E+00	11067964	NT	Homo sapiens KIAA0628 gene product (KIAA0628), mRNA
11064	24027		3.83	1.1E+00	AF068942.1	NT	Klebsiella pneumoniae cytochrome c oxidase subunit 2 (cox2) gene, mitochondrial gene encoding mitochondrial protein, partial cds

Page 27 of 546
Table 4
Single Exon Probes Expressed in Bone Marrow

Probe SEQ ID NO:	Exon SEQ ID NO:	ORF SEQ ID NO:	Expression Signal	Most Similar (Top) Hit BLAST E Value	Top Hit Accession No.	Top Hit Database Source	Top Hit Descriptor
11435	18340		4.65	1.1E+00	8922873	NT	Homo sapiens hypothetical protein FLJ11280 (FLJ11280), mRNA
11439	24382	37921	3.12	1.1E+00	AF012862.1	NT	Petrosselinum crispum cytosolic glucose-6-phosphate dehydrogenase 1 (cG6PDH1) mRNA, complete cds
11439	24382	37922	3.12	1.1E+00	AF012862.1	NT	Petrosselinum crispum cytosolic glucose-6-phosphate dehydrogenase 1 (cG6PDH1) mRNA, complete cds
11683	24649	38227	3.44	1.1E+00	AI809699.1	EST_HUMAN	wf76a11.x1 Soares_NFL_I_GBC_S1 Homo sapiens cDNA clone IMAGE:2361648 3'
12439	25202		3.12	1.1E+00	P07868	SWISSPROT	LOW TEMPERATURE ESSENTIAL PROTEIN
12539	25263	31807	1.55	1.1E+00	AF216696.1	NT	Teenia solium immunogenic protein Ts78 mRNA, partial cds
12661	25796		2.26	1.1E+00	AF234189.1	NT	Dictyostelium discoideum isopentenyl pyrophosphate isomerase (Dipi) mRNA, complete cds
88	13215		1.55	1.0E+00	U23808.1	NT	Xenopus laevis rhodopsin gene, complete cds
113	13224	26148	1.86	1.0E+00	D88425.1	NT	Cavia cobaya mRNA for serine/threonine kinase, complete cds
418	13491		3.03	1.0E+00	AB021694.1	NT	Marchantia polymorpha genes for 26S rRNA, 5S rRNA, 16S rRNA, 5.8S rRNA and 26S rRNA
578	13647	26560	2.22	1.0E+00	AJ251690.1	NT	Giardia tigrina mRNA for homeodomain transcription factor (so gene)
678	13741	26698	6.74	1.0E+00	AL163218.2	NT	Homo sapiens chromosome 21 segment HS21C018
678	13742		1.35	1.0E+00	AF125984.1	NT	Aedes aegypti mucin-like protein MUC1 mRNA, complete cds
1386	15866		3.53	1.0E+00	X80418.1	NT	V. carteri Algal-CAM mRNA
1771	14900	27796	0.97	1.0E+00	AB006531.1	NT	Plautia stali intestine virus RNA for nonstructural polyprotein, capsid protein precursor, complete cds
2493	15498	28521	1.42	1.0E+00	P48355	SWISSPROT	DNA GYRASE SUBUNIT B
2493	15498	28522	1.42	1.0E+00	P48355	SWISSPROT	DNA GYRASE SUBUNIT B
2887	15946	28861	4.42	1.0E+00	P24008	SWISSPROT	3-OXO-5-ALPHA-STEROID 4-DEHYDROGENASE 1 (STEROID 5-ALPHA-REDUCTASE 1) (SR TYPE 1)
2887	15946	28862	4.42	1.0E+00	P24008	SWISSPROT	3-OXO-5-ALPHA-STEROID 4-DEHYDROGENASE 1 (STEROID 5-ALPHA-REDUCTASE 1) (SR TYPE 1)
2978	16036		0.81	1.0E+00	O14226	SWISSPROT	HYPOTHETICAL 67.9 KD PROTEIN C8F12.08C IN CHROMOSOME I
3212	16267	29190	1.08	1.0E+00	AA628453.1	EST_HUMAN	a28g08.at Soares_tetus_Nib2HF8_9w Homo sapiens cDNA clone IMAGE:1032850 3' similar to
3613	13215		0.93	1.0E+00	U23808.1	NT	WP-C42D8.3 C504204 ;contains element MER22 MER22 repetitive element ;
3693	16736	29649	1.33	1.0E+00	AJ223616.1	NT	Xenopus laevis rhodopsin gene, complete cds
						NT	Agaricus bisporus mRNA for tyrosinase
4097	17131	30024	1.16	1.0E+00	AF223391.1	NT	Homo sapiens calcium channel alpha1E subunit (CACNA1E) gene, exons 7-49, and partial cds, alternatively spliced
4304	17333		0.73	1.0E+00	8922245	NT	Homo sapiens hypothetical protein FLJ10139 (FLJ10139), mRNA
4843	17860	30755	1.54	1.0E+00	AL163247.2	NT	Homo sapiens chromosome 21 segment HS21C047
5053	18065		0.9	1.0E+00	D10852.1	NT	Rattus norvegicus mRNA for N-acetylglucosaminyltransferase III, complete cds

Page 28 of 546
Table 4

Single Exon Probes Expressed in Bone Marrow

Probe SEQ ID NO:	Exon SEQ ID NO:	ORF SEQ ID NO:	Expression Signal	Most Similar (Top) Hit BLAST E Value	Top Hit Accession No.	Top Hit Database Source	Top Hit Descriptor
5354	18469	31328	3.23	1.0E+00	Z97022.1	NT	Hordeum vulgare gene encoding cysteine proteinase
5949	19035	32228	5.05	1.0E+00	AF248054.1	NT	Bos taurus microsomal calcium activated neutral protease 1 (CAPN1) gene, exons 11-20, and partial cds
5949	19035	32229	5.05	1.0E+00	AF248054.1	NT	Bos taurus microsomal calcium activated neutral protease 1 (CAPN1) gene, exons 11-20, and partial cds
6067	19148	32360	1.44	1.0E+00	Z97341.2	NT	Arabidopsis thaliana DNA chromosome 4, ESSA1 FCA contig fragment No. 6
6236	19309	32541	4.52	1.0E+00	P04501	SWISSPROT	FIBER PROTEIN
6243	19316	32546	1.77	1.0E+00	AW452782.1	EST_HUMAN	UI-HB13-ab-4-09-0-U1.s1 NCJ_CGAP_Sub5 Homo sapiens cDNA clone IMAGE:3068969 3'
6637	19695	32972	2.12	1.0E+00	U75902.1	NT	Mus musculus subtilisin-like serine protease LPC (PC7) gene, exons 1 to 9, partial cds
6691	19748	33025	0.68	1.0E+00	AF104689.1	NT	Homo sapiens cell cycle protein (PA2G4) gene, exons 2 through 5
6786	19841		0.98	1.0E+00	P46508	SWISSPROT	SRB-11 PROTEIN
6813	19867	33155	0.71	1.0E+00	BE797718.1	EST_HUMAN	601581891F1 NIH_MGC_7 Homo sapiens cDNA clone IMAGE:3836382 5'
6813	19867	33156	0.71	1.0E+00	BE797718.1	EST_HUMAN	601581891F1 NIH_MGC_7 Homo sapiens cDNA clone IMAGE:3836382 5'
6844	20168	33491	1.17	1.0E+00	Y11204.1	NT	V. carteri gene encoding vacuolysin
7033	18365	31252	0.64	1.0E+00	U63721.1	NT	Human elastin (ELN) gene, partial cds, and LIM-kinase (LIMK1) gene, complete cds
7345	20316	33661	1	1.0E+00	S52770.1	NT	Insulin-like growth factor-binding protein 4 [catilla, pulmonary artery endothelial cells, mRNA, 2028 nt]
7719	20678		8.7	1.0E+00	P20273	SWISSPROT	B-CELL RECEPTOR CD22 PRECURSOR (LEU-14) (B-LYMPHOCYTE CELL ADHESION MOLECULE) (BL-CAM)
7979	20918	34308	1.52	1.0E+00	AF182531.1	NT	Homo sapiens endothelin-converting enzyme 2 (ECE2) mRNA, complete cds
7997	20939	34330	6.07	1.0E+00	AA775191.1	EST_HUMAN	ac79b08.s1 Strategene lung (#637210) Homo sapiens cDNA clone IMAGE:868781 3'
8165	21103		0.57	1.0E+00	BF078213.1	EST_HUMAN	602153782F1 NIH_MGC_83 Homo sapiens cDNA clone IMAGE:4294727 5'
8294	21263	34673	1.49	1.0E+00	BE868267.1	EST_HUMAN	601443950F1 NIH_MGC_65 Homo sapiens cDNA clone IMAGE:3848005 5'
8294	21263	34674	1.49	1.0E+00	BE868267.1	EST_HUMAN	601443950F1 NIH_MGC_65 Homo sapiens cDNA clone IMAGE:3848005 5'
8481	18065		1.22	1.0E+00	D10852.1	NT	Rattus norvegicus mRNA for N-acetylglucosaminyltransferase III, complete cds
							PEROXISOMAL HYDRATASE-DEHYDROGENASE-EPIMERASE (HDE) (MULTIFUNCTIONAL BETA-OXIDATION PROTEIN) (MFP) [INCLUDES: 2-ENOYL-COA HYDRATASE; D-3-HYDROXYACYL COA DEHYDROGENASE]
8693	21661	35084	2.31	1.0E+00	Q02207	SWISSPROT	PEROXISOMAL HYDRATASE-DEHYDROGENASE-EPIMERASE (HDE) (MULTIFUNCTIONAL BETA-OXIDATION PROTEIN) (MFP) [INCLUDES: 2-ENOYL-COA HYDRATASE; D-3-HYDROXYACYL COA DEHYDROGENASE]
8693	21661	35085	2.31	1.0E+00	Q02207	SWISSPROT	PEROXISOMAL HYDRATASE-DEHYDROGENASE-EPIMERASE (HDE) (MULTIFUNCTIONAL BETA-OXIDATION PROTEIN) (MFP) [INCLUDES: 2-ENOYL-COA HYDRATASE; D-3-HYDROXYACYL COA DEHYDROGENASE]
8821	21788		0.83	1.0E+00	P51784	SWISSPROT	UBIQUITIN CARBOXYL-TERMINAL HYDROLASE 11 (UBIQUITIN THIOLESTERASE 11) (UBIQUITIN-SPECIFIC PROCESSING PROTEASE 11) (DEUBIQUITINATING ENZYME 11)

Page 29 of 546
Table 4
Single Exon Probes Expressed in Bone Marrow

Probe SEQ ID NO:	Exon SEQ ID NO:	ORF SEQ ID NO:	Expression Signal	Most Similar (Top) Hit BLAST E Value	Top Hit Accession No.	Top Hit Database Source	Top Hit Descriptor
8855	21822	35242	0.44	1.0E+00	Q8Y5T5	SWISSPROT	UBIQUITIN CARBOXYL-TERMINAL HYDROLASE 16 (UBIQUITIN THIOLESTERASE 16) (UBIQUITIN-SPECIFIC PROCESSING PROTEASE 16) (DEUBIQUITINATING ENZYME 16) (UBIQUITIN PROCESSING PROTEASE UBP-M)
8855	21822	35243	0.44	1.0E+00	Q8Y5T5	SWISSPROT	UBIQUITIN CARBOXYL-TERMINAL HYDROLASE 16 (UBIQUITIN THIOLESTERASE 16) (UBIQUITIN-SPECIFIC PROCESSING PROTEASE 16) (DEUBIQUITINATING ENZYME 16) (UBIQUITIN PROCESSING PROTEASE UBP-M)
8883	25698		2.37	1.0E+00	BE147331.1	EST_HUMAN	RC1-HT0229-181089-011-e08 HT0229 Homo sapiens cDNA
8923	21889	35316	0.89	1.0E+00	U42720.2	NT	SriLan Immunodeficiency virus Gag protein (gag) gene, complete cds; Pcl protein (pcl) gene, partial cds; and Vif protein (vif), Vpr protein (vpr), Tat protein (tat), Rev protein (rev), Vpu protein (vpu), Env protein (env), and Nef protein (nef) genes, >
9076	22041	35484	1.55	1.0E+00	M38427.1	NT	Human Immunodeficiency virus type 1 (HIV-1), isolate SF33,
9826	22570	36019	2.14	1.0E+00	BE907692.1	EST_HUMAN	801497681F1 NIH_MGC_70 Homo sapiens cDNA clone IMAGE:3889421 5'
9836	22772	36227	1.25	1.0E+00	8753429	NT	Mus musculus chloride channel calcium activated 1 (Cicac1), mRNA
9836	22772	36228	1.25	1.0E+00	8753429	NT	Mus musculus chloride channel calcium activated 1 (Cicac1), mRNA
9869	22898	36359	1.94	1.0E+00	AV689564.1	EST_HUMAN	AV689554 GKC Homo sapiens cDNA clone GKCCYA11 5'
9874	22901	36384	1.23	1.0E+00	U44982.1	NT	Xenopus laevis zona pellucida C glycoprotein precursor (XZPC) mRNA, complete cds
9874	22901	36385	1.23	1.0E+00	U44982.1	NT	Xenopus laevis zona pellucida C glycoprotein precursor (XZPC) mRNA, complete cds
10212	23137	36824	0.56	1.0E+00	X15498.1	NT	Human Coronavirus gene for membrane protein
10212	23137	36825	0.56	1.0E+00	X15498.1	NT	Human Coronavirus gene for membrane protein
10471	23393	36889	0.72	1.0E+00	5174562	NT	Homo sapiens MHC binding factor, beta (MHC8FB) mRNA
10471	23393	36890	0.72	1.0E+00	5174562	NT	Homo sapiens MHC binding factor, beta (MHC8FB) mRNA
10564	23486	36980	0.63	1.0E+00	A1077920.1	EST_HUMAN	cy15d07.a1 Soares_sarcomat_fibroblasts_NH-SF Homo sapiens cDNA clone IMAGE:1065901 3'
10690	23612	37106	3.7	1.0E+00	AV758825.1	EST_HUMAN	AV758825 BM Homo sapiens cDNA clone BMFAW04 5'
10842	23762	37262	20.08	1.0E+00	AA004982.1	EST_HUMAN	z94a02.r1 Soares_fetal_liver_spleen_1NFLS_S1 Homo sapiens cDNA clone IMAGE:428908 5'
10842	23762	37263	20.08	1.0E+00	AA004982.1	EST_HUMAN	z94a02.r1 Soares_fetal_liver_spleen_1NFLS_S1 Homo sapiens cDNA clone IMAGE:428908 5'
10876	23798	37297	1.18	1.0E+00	L11910.1	NT	Human retinoblastoma susceptibility gene exons 1-27, complete cds
12046	19459	31328	1.66	1.0E+00	Z97022.1	NT	Hardycore vulgaris gene encoding cysteine protease
12327	25129		3.26	1.0E+00	P15306	SWISSPROT	THROMBOMODULIN PRECURSOR (FETOMODULIN) (TM)
12650	25333		2.49	1.0E+00	AW976184.1	EST_HUMAN	EST388293 MAGE resequences, MAGN Homo sapiens cDNA
1575	14608	27580	3.22	9.9E-01	AF245455.1	NT	Drosophila melanogaster regulator of G-protein signalling LOCO III mRNA, complete cds
1575	14608	27581	3.22	9.9E-01	AF245455.1	NT	Drosophila melanogaster regulator of G-protein signalling LOCO III mRNA, complete cds
2845	15642	28686	1	9.9E-01	AL163302.2	NT	Homo sapiens chromosome 21 segment HS21C102
3619	16663		1.1	9.9E-01	AF174585.1	NT	Apple mosaic virus RNA 2 putative polymerase gene, complete cds
5717	18811	31890	8.62	9.9E-01	P49657	SWISSPROT	SERINE/THREONINE PROTEIN KINASE MINIBRAIN

Page 30 of 546
Table 4
Single Exon Probes Expressed in Bone Marrow

Probe SEQ ID NO:	Exon SEQ ID NO:	ORF SEQ ID NO:	Expression Signal	Most Similar (Top) Hit BLAST E Value	Top Hit Accession No.	Top Hit Database Source	Top Hit Descriptor
5989	19064	32264	0.79	9.8E-01	Q08632	SWISSPROT	PROBABLE OXIDOREDUCTASE ZK1290.5 IN CHROMOSOME II
9816	22560		1.4	9.8E-01	U66667.1	NT	Lycopodium esculentum putative Mt1 copy 1 nematode-resistance gene
9913	22734		3.02	9.8E-01	Q28642	SWISSPROT	B2 BRADYKININ RECEPTOR (BK-2 RECEPTOR)
11069	24032	37566	1.48	9.8E-01	AJ005029.1	NT	Danio rerio mRNA for Ebf-like receptor tyrosine kinase rtk8
524	13595	26513	1.12	9.8E-01	P22567	SWISSPROT	AMINO-ACID ACETYLTRANSFERASE (N-ACETYL-GLUTAMATE SYNTHASE) (AGS) (NAGS)
2305	15317		1.28	9.8E-01	AJ003108.1	NT	Callitriche jacobus UBE1 gene derived retroposon on the Y chromosome
2813	15805		1.29	9.8E-01	AF174644.1	NT	Xenopus laevis rac GTPase mRNA, complete cds
7406	20374	33725	4.12	9.8E-01	AJ302158.1	NT	Enterobacteriaceae sp. JM983 partial groES gene for GroES-like protein and partial groEL gene for GroEL-like protein, isolate JM983
7406	20374	33726	4.12	9.8E-01	AJ302158.1	NT	Enterobacteriaceae sp. JM983 partial groES gene for GroES-like protein and partial groEL gene for GroEL-like protein, isolate JM983
7907	20850	34236	0.89	9.8E-01	BF034016.1	EST_HUMAN	601456337F1 NIH_MGC_68 Homo sapiens cDNA clone IMAGE:3860049 5'
7907	20850	34237	0.89	9.8E-01	BF034016.1	EST_HUMAN	601456337F1 NIH_MGC_68 Homo sapiens cDNA clone IMAGE:3860049 5'
9069	22035	35458	0.88	9.8E-01	P38662	SWISSPROT	PHOSPHOGLUCOMUTASE (GLUCOSE PHOSPHOMUTASE) (PGM)
10804	23725		0.53	9.8E-01	AA825565.1	EST_HUMAN	cd55404.s1 NCJ_CGAP_GCB1 Homo sapiens cDNA clone IMAGE:1371847 3'
11339	24289	37813	2.06	9.8E-01	BE258705.1	EST_HUMAN	601110258F1 NIH_MGC_16 Homo sapiens cDNA clone IMAGE:3350750 5'
11339	24289	37814	2.06	9.8E-01	BE258705.1	EST_HUMAN	601110258F1 NIH_MGC_16 Homo sapiens cDNA clone IMAGE:3350750 5'
12645	25268		1.41	9.8E-01	U52111.2	NT	Homo sapiens X28 region near ALD locus containing dual specificity phosphatase 9 (DUSP9), ribosomal protein L18a (RPL18a), Cez2/Calmodulin-dependent protein kinase 1 (CAMK1), creatine transporter (CRT), CDM protein (CDM), adrenoleukodystrophy protein >
7366	20336	33688	2.3	9.7E-01	U28716.1	NT	Drosophila melanogaster sodium channel protein (para) gene, exons 8,10,11,12 and optional segments b, c, d and e, partial cds
8848	21815	35235	1.81	9.7E-01	AF149112.1	NT	Triticum aestivum stripe rust resistance protein Yr10 (Yr10) gene, complete cds
8854	21821	35241	1.33	9.7E-01	M90544.1	NT	Salmonella typhimurium adenine-methyltransferase (mod) and restriction endonuclease (res)
11505	24447		3.87	9.7E-01	BF611209.1	EST_HUMAN	U1-H-B14-ec-e-07-Q-U1.s1 NCJ_CGAP_Sub8 Homo sapiens cDNA clone IMAGE:3085140 3'
12117	24687	36589	2.29	9.7E-01	U87514.1	NT	Dictyostelium discoideum CAR3 gene, promoter region
4486	17511	30398	1.68	9.6E-01	AW78674.1	EST_HUMAN	PM2-UM0053-240300-005-t12 UM0053 Homo sapiens cDNA
5179	18188	31065	0.9	9.6E-01	7862375	NT	Homo sapiens KIAA00914 gene product (KIAA00914), mRNA
5947	18937	32121	3.85	9.6E-01	Z70556.1	NT	Parvovirus B19 DNA, patient C, genome position 2448-2994
5947	18937	32122	3.85	9.6E-01	Z70556.1	NT	Parvovirus B19 DNA, patient C, genome position 2448-2994
6910	18932	33258	0.67	9.6E-01	Z97341.2	NT	Arabisopsis thaliana DNA chromosome 4, ESSA I FCA contig fragment No. 6
8735	21703		1.21	9.6E-01	X95275.1	NT	P. falciparum complete gene map of plastic-like DNA (IR-A)
9203	22169	35569	0.51	9.6E-01	L81138.1	NT	Rattus norvegicus (strain R21) Rps27 gene, complete cds

Page 31 of 546
Table 4

Single Exon Probes Expressed in Bone Marrow

Probe SEQ ID NO:	Exon SEQ ID NO:	ORF SEQ ID NO:	Expression Signal	Most Similar (Top) Hit BLAST E Value	Top Hit Accession No.	Top Hit Database Source	Top Hit Descriptor
9446	22410	35947	0.44	9.8E-01	AF228843.1	NT	Mus musculus WNT-2 gene, partial cds; putative ankyrin-related protein and cystic fibrosis transmembrane conductance regulator (CFTR) genes, section 1 of 2 of the complete cds; and unknown gene
11848	24731	38317	3.04	9.8E-01	AV752805.1	EST_HUMAN	AV752805 NP2 Homo sapiens cDNA clone NPDBAG06 5'
11848	24731	38318	3.04	9.8E-01	AV752805.1	EST_HUMAN	AV752805 NP2 Homo sapiens cDNA clone NPDBAG06 5'
12223	25060		2.19	9.8E-01	11421722	NT	Homo sapiens centrosomal protein 2 (CEP2), mRNA
12839	25874	31416	3.03	9.8E-01	U91423.1	NT	Sphynx fibro NADH dehydrogenase subunit 2 (NADH2) gene, mitochondrial gene encoding mitochondrial protein, partial cds
2484	15488	28512	1.03	9.8E-01	7705591	NT	Homo sapiens CGI-125 protein (LOC51003), mRNA
3788	16836	29741	2.39	9.8E-01	BE902340.1	EST_HUMAN	601675639F1 NIH_MGC_21 Homo sapiens cDNA clone IMAGE:3958473 5'
3788	16836	29742	2.39	9.8E-01	BE902340.1	EST_HUMAN	601675639F1 NIH_MGC_21 Homo sapiens cDNA clone IMAGE:3958473 5'
8354	22319	35745	0.68	9.8E-01	AI190162.1	EST_HUMAN	qd57407.x1 Soares_testis_NHT Homo sapiens cDNA clone IMAGE:1733581 3'
9461	22425	35863	1.1	9.8E-01	AW681102.1	EST_HUMAN	RC1-CT0285-241199-011-302 CT0285 Homo sapiens cDNA
11674	24512	38088	1.5	9.8E-01	BF218771.1	EST_HUMAN	601885103F1 NIH_MGC_57 Homo sapiens cDNA clone IMAGE:4103630 5'
11780	23935	37456	1.52	9.8E-01	AW283798.1	EST_HUMAN	UI-H-B12-ahp-f03-0-UI.s1 NC1_CGAP_Sub4 Homo sapiens cDNA clone IMAGE:2127677 3'
3214	16289		3.77	9.4E-01	AF169990.1	NT	Baronella clarridgeae RNA polymerase beta subunit (rpoB) gene, partial cds
3231	16286		1.93	9.4E-01	AF169990.1	NT	Pimphelia brachycarpa zinc finger protein (ZFP1) mRNA, complete cds
9217	22183	35616	0.69	9.4E-01	M90724.1	NT	Human Fo-gamma-receptorIIA (FCGR2A) gene, exon 4
12490	25235		2.09	9.4E-01	BE781251.1	EST_HUMAN	601466703F1 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:3869029 5'
12638	25780		1.93	9.4E-01	11419857	NT	Homo sapiens epidermal growth factor receptor (avian erythroblastic leukemia viral (v-erb-b) oncogene homolog) (EGFR), mRNA
1745	14774		1.34	9.3E-01	AF242382.1	NT	Homo sapiens phytyl-CoA hydroxylase (PHYH) gene, exon 5
2641	15639	28663	1.01	9.3E-01	BE071172.1	EST_HUMAN	RC5-BT0503-271199-011-B01 BT0503 Homo sapiens cDNA
4098	17102	28993	0.82	9.3E-01	M20219.1	NT	Bovine papillomavirus type 2, complete genome
4098	17102	28994	0.82	9.3E-01	M20219.1	NT	Bovine papillomavirus type 2, complete genome
5673	18788	31940	1.47	9.3E-01	AF213884.1	NT	Homo sapiens nuclear factor of kappa light polypeptide gene enhancer in B-cells 1 (NFKB1) gene, complete cds
5761	18854	32034	3.92	9.3E-01	L36189.1	NT	Spodoptera frugiperda methylenetetrahydrofolate dehydrogenase mRNA, complete cds
7554	20517		0.76	9.3E-01	AF270648.1	NT	Plasmodium falciparum mature parasite-infected erythrocyte surface antigen (MESA) gene, complete cds
8402	21371	34779	1.75	9.3E-01	AA847040.1	EST_HUMAN	oe09b03.s1 NC1_CGAP_OV2 Homo sapiens cDNA clone IMAGE:1385357
9165	22131		0.95	9.3E-01	AF061861.1	NT	Xenopus laevis COOH zinc finger protein C3H-2 (C3H-2) mRNA, complete cds
9259	22255	35685	0.91	9.3E-01	AL161534.2	NT	Arabidopsis thaliana DNA chromosome 4, contig fragment No. 34
12853	25528	31713	1.34	9.3E-01	11440298	NT	Homo sapiens inositol 1,4,5-trisphosphate receptor, type 2 (ITPR2), mRNA

Page 32 of 546
Table 4

Single Exon Probes Expressed in Bone Marrow

Probe SEQ ID NO:	Exon SEQ ID NO:	ORF SEQ ID NO:	Expression Signal	Most Similar (Top) Hit BLAST E Value	Top Hit Accession No.	Top Hit Database Source	Top Hit Descriptor
12965	25534		3.29	9.3E-01	AF271207.1	NT	Aedes triseriatus putative large subunit ribosomal protein rPL34 mRNA, complete cds
3253	16307	29231	3.14	9.2E-01	BE622702.1	EST_HUMAN	601441338T1 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:3916184 3'
5806	18698		1.73	9.2E-01	7106410	NT	Mus musculus solute carrier family 30 (zinc transporter), member 4 (SLC30A4), mRNA
6101	19180	32399	4.04	9.2E-01	BF037586.1	EST_HUMAN	601461153F1 NIH_MGC_68 Homo sapiens cDNA clone IMAGE:3864661 5'
6789	19843	33127	0.51	9.2E-01	MG4703.1	NT	N. crassa vely-RNA synthetase (cyl-204un-3) gene
10018	22945	39412	0.77	9.2E-01	AL161565.2	NT	Arabidopsis thaliana DNA chromosome 4, contig fragment No. 65
10105	23031	36509	1.21	9.2E-01	6871677	NT	Mus musculus carbonic anhydrase 4 (Car4), mRNA
10628	23550	37050	3.42	9.2E-01	11430963	NT	Homo sapiens lysosomal apyrase-like protein 1 (LALP1), mRNA
10780	23701	37199	1.84	9.2E-01	BF593251.1	EST_HUMAN	7658408.x1 NC1_CGAP_Kid11 Homo sapiens cDNA clone IMAGE:3578219 3' similar to SW:NU5M_TRYBB
12031	24907	38501	1.54	9.2E-01	BF132402.1	EST_HUMAN	P04540 NADH-LUBIQUINONE OXIDOREDUCTASE CHAIN 5;
1629	14982	27638	2.31	9.1E-01	T36875.1	EST_HUMAN	601820312F1 NIH_MGC_58 Homo sapiens cDNA clone IMAGE:4052018 5'
2134	15151		2.08	9.1E-01	8223056	NT	ye52f01.s1 Soares fetal liver spleen 1NFLS Homo sapiens cDNA clone IMAGE:121369 3' similar to contains
2582	15583	25602	1.12	9.1E-01	AF062919.1	NT	Alu repetitive element
3218	16273	29195	1.11	9.1E-01	T26418.1	EST_HUMAN	Homo sapiens hypothetical protein FLJ20048 (FLJ20048), mRNA
3218	16279	29196	1.11	9.1E-01	T26418.1	EST_HUMAN	Pseudomonas fluorescens DNA polymerase III (dnaE) gene, complete cds
6291	19363	32602	1.68	9.1E-01	L36033.1	NT	AB200G8R Infant brain, LLNL array of Dr. M. Soares 1NIB Homo sapiens cDNA clone LLAB200G8 5'
6655	19712	32989	2.94	9.1E-01	Q61704	SWISSPROT	AB200G8R Infant brain, LLNL array of Dr. M. Soares 1NIB Homo sapiens cDNA clone LLAB200G8 5'
7827	20775	34152	18.4	9.1E-01	AA806623.1	EST_HUMAN	Human pre-B cell stimulating factor homologue (SDF1b) mRNA, complete cds
8011	20949	34342	2.58	9.1E-01	U72965.1	NT	INTER-ALPHA-TRYPsin INHIBITOR HEAVY CHAIN H3 PRECURSOR (ITI HEAVY CHAIN H3)
10536	23458	36955	0.45	9.1E-01	P38432	SWISSPROT	cb71g08.s1 NC1_CGAP_GCB1 Homo sapiens cDNA clone IMAGE:1336862 3'
12580	25867		10.31	9.1E-01	AF050113.1	NT	Rattus norvegicus Rab3 GDP/GTP exchange protein mRNA, complete cds
4408	17494	30319	1.77	9.0E-01	AF098810.1	NT	P80-COLLIN
7622	20582	33946	0.65	9.0E-01	L42547.1	NT	Homo sapiens uncoupling protein-3 (UCP3) gene, complete cds
7652	20612		1.32	9.0E-01	D38621.1	NT	Homo sapiens neuroxin III-alpha gene, partial cds
9704	22867	36112	0.55	9.0E-01	AF086761.1	NT	Danio rerio LIM class homeodomain protein (lim5) mRNA, complete cds
10189	23114	36598	0.44	9.0E-01	U39702.1	NT	Xenopus laevis gene for aldolase, complete cds
							Danio rerio semaphorin Z1a mRNA, complete cds
							Danio rerio semaphorin Z1a mRNA, complete cds
							Mycoplasma genitalium section 24 of 51 of the complete genome
5781	18873	32054	2.37	8.9E-01	AF026198.1	NT	Fugu rubripes neural cell adhesion molecule L1 homolog (L1-CAM) gene, complete cds; putative protein 1 (PUT1) gene, partial cds; mitosis-specific chromosome segregation protein SMC1 homolog (SMC1) gene, complete cds; and calcium channel alpha-1 subunit?
6377	19445		1.28	8.9E-01	X60986.1	NT	Rabbit MHC fragment RLA-DF DNA

Table 4

Single Exon Probes Expressed in Bone Marrow

Probe SEQ ID NO:	Exon SEQ ID NO:	ORF SEQ ID NO:	Expression Signal	Most Similar (Top) Hit BLAST E Value	Top Hit Accession No.	Top Hit Database Source	Top Hit Descriptor
6605	25663	32937	0.69	8.9E-01	BF217939.1	EST_HUMAN	601882708F1 NIH_MGC 57 Homo sapiens cDNA clone IMAGE:4095216 5'
6605	25663	32938	0.69	8.9E-01	BF217939.1	EST_HUMAN	601882708F1 NIH_MGC 57 Homo sapiens cDNA clone IMAGE:4095216 5'
7408	20375		0.54	8.9E-01	AB042287.1	NT	Homo sapiens PTS gene for 6-pyruvyltetrahydropterin synthase, complete cds
8653	21521		0.43	8.9E-01	AF280225.1	NT	Homo sapiens TESTIN 2 and TESTIN 3 genes, complete cds, alternatively spliced
8769	21736	35157	1.09	8.9E-01	AF259867.1	NT	Orithona nana cytochrome-c oxidase subunit 1 (cod) gene, partial cds; mitochondrial gene for mitochondrial product
12074	24948	38541	2.99	8.9E-01	AE003944.1	NT	Xylella fastidiosa, section 90 of 229 of the complete genome
12420	25190		5.46	8.9E-01	AE002186.2	NT	Chlamydia pneumoniae AFR39, section 21 of 94 of the complete genome
4573	17595	30489	1.56	8.8E-01	O28350	SWISSPROT	PUTATIVE F420-DEPENDENT NADP REDUCTASE
5238	18246	31118	0.9	8.8E-01	L41854.1	NT	Trypanosoma brucei microtubule-associated protein (MAP15) mRNA, 3' end of cds
6447	18549	31462	0.7	8.8E-01	AF310617.1	NT	Pseudorabies virus Ea glycoprotein M gene, complete cds
11418	24362	37897	2.31	8.8E-01	Z28337.1	NT	M. aeruginosa (HUB 5-2-4) DNA from plasmid PMA1
12082	24954	38549	6.51	8.8E-01	AA808055.1	EST_HUMAN	cc38111.1 NCI CGAP_G031 Homo sapiens cDNA clone IMAGE:1352037 3' similar to contains Alu repetitive element; contains element MER22 repetitive element;
12237	25952		2.3	8.8E-01	D90911.1	NT	Synechococcus sp. POC9803 complete genome, 13/27, 1576593-1719843
465	13538	28484	1.78	8.7E-01	AF109953.2	NT	Homo sapiens SOS1 (SOS1) gene, partial cds
2411	15418	28442	0.91	8.7E-01	5901893	NT	Homo sapiens AT-binding transcription factor 1 (ATBF1), mRNA
2885	15944	28859	5.42	8.7E-01	AA695663.1	EST_HUMAN	nm05511.1 NCI CGAP_P4.1 Homo sapiens cDNA clone IMAGE:1076877
							Pseudomonas aeruginosa lipdimerase (top), putative transcriptional regulatory protein OhbR (ohbR), ortho-halobenzoate 1,2-dioxygenase beta-ISP protein OhbA (ohbA), OhbC (ohbC), ortho-halobenzoate 1,2-dioxygenase alpha-ISP protein OhbB (ohbB), and put
5043	18059		2.78	8.7E-01	AF121970.1	NT	601883175F1 NIH_MGC 57 Homo sapiens cDNA clone IMAGE:4095378 5'
5247	18255	31125	0.97	8.7E-01	BF219308.1	EST_HUMAN	RC4-NN0057-120500-019-c07 NN0057 Homo sapiens cDNA
8374	21343	34764	0.65	8.7E-01	AW897335.1	EST_HUMAN	qt38606.x1 Soares NFL_T_GBC S1 Homo sapiens cDNA clone IMAGE:1846786 3'
8282	22248	35677	0.71	8.7E-01	AL239456.1	EST_HUMAN	qt38606.x1 Soares NFL_T_GBC S1 Homo sapiens cDNA clone IMAGE:1846786 3'
9282	22248	35678	0.71	8.7E-01	AL239456.1	EST_HUMAN	qt38606.x1 Soares NFL_T_GBC S1 Homo sapiens cDNA clone IMAGE:1846786 3'
10095	23021	36498	1.32	8.7E-01	AE004963.1	NT	Pseudomonas aeruginosa PA01, section 524 of 529 of the complete genome
11181	24197	37689	4.56	8.7E-01	BF363970.1	EST_HUMAN	QV0-NN1021-100800-337-c03 NN1021 Homo sapiens cDNA
12042	24917	38512	3.97	8.7E-01	BF107694.1	EST_HUMAN	601823684R1 NIH_MGC 79 Homo sapiens cDNA clone IMAGE:4043564 3'
12042	24917	38513	3.97	8.7E-01	BF107694.1	EST_HUMAN	601823684R1 NIH_MGC 79 Homo sapiens cDNA clone IMAGE:4043564 3'
475	13547		2.65	8.6E-01	X17012.1	NT	Rat IGF1 gene for insulin-like growth factor II
859	13915	28874	4.32	8.6E-01	W60089.1	EST_HUMAN	z044e03 r1 Soares_fetal_heart_NbHH19W Homo sapiens cDNA clone IMAGE:343516 5'
							Homo sapiens cytochrome P450, subfamily XXVIIA (steroid 27-hydroxylase, cerebrotendinous xanthomatosis), polypeptide 1 (CYP27A1b) mRNA
2278	15291	28316	1	8.6E-01	4503210	NT	Arabidopsis thaliana DNA chromosome 4, contig fragment No. 65
3636	18678	28592	0.87	8.6E-01	AL161555.2	NT	

Page 34 of 546
Table 4
Single Exon Probes Expressed in Bone Marrow

Probe SEQ ID NO:	Exon SEQ ID NO:	ORF SEQ ID NO:	Expression Signal	Most Similar (Top) Hit BLAST E Value	Top Hit Accession No.	Top Hit Database Source	Top Hit Descriptor
3813	18853	28760	1.38	8.6E-01	U49724.1	NT	Drosophila melanogaster marlin (Dmarlin) mRNA, complete cds
5207	18216	31081	2.6	8.6E-01	BE147809.1	EST_HUMAN	RC1-HT0229-180300-019-c05 HT0229 Homo sapiens cDNA
6001	18084	32283	7.78	8.6E-01	X60547.1	NT	Chicken lipoprotein lipase gene
6001	18084	32284	7.78	8.6E-01	X60547.1	NT	Chicken lipoprotein lipase gene
6515	25661	32635	0.54	8.6E-01	S76772.1	NT	polyprotein [Coxsackie B4 virus CB4, host=rice, E2, originally derived from Edwards CB4 human strain, Genomic RNA Complete, 7397 nt]
6867	18920	33215	1.7	8.6E-01	AF143732.1	NT	Grus carolinensis recombination activating protein 1 (RAG-1) gene, partial cds
6867	18920	33216	1.7	8.6E-01	AF143732.1	NT	Grus carolinensis recombination activating protein 1 (RAG-1) gene, partial cds
7770	20723		0.81	8.6E-01	AE000591.1	NT	Helicobacter pylori 26695 section 69 of 134 of the complete genome
8260	21228		1.29	8.6E-01	AP001818.1	NT	Bacillus halodurans genomic DNA, section 12/14
8377	21346	34767	0.51	8.6E-01	AF077837.1	NT	Drosophila melanogaster collapse response mediator protein (CRMP) mRNA, complete cds
10044	22971		0.46	8.6E-01	AE000979.1	NT	Archaeoglobus fulgidus section 128 of 172 of the complete genome
12798	25717		1.44	8.6E-01	AL112162.1	NT	Botrytis cinerea strain T4 cDNA library under conditions of nitrogen deprivation
6888	18940	33235	1.5	8.5E-01	AF165214.1	NT	Bacteriophage D3, complete genome
7768	20721	34083	2.49	8.5E-01	BE542812.1	EST_HUMAN	601067107F1 NIH_MGC_10 Homo sapiens cDNA clone IMAGE:3453505 5'
8323	21282	34708	0.42	8.5E-01	AL161572.2	NT	Arabidopsis thaliana DNA chromosome 4, contig fragment No. 68
8761	21728	35150	0.93	8.5E-01	P06801	SWISSPROT	SEGMENTATION PROTEIN PAIRED
8761	21728	35151	0.93	8.5E-01	P06801	SWISSPROT	SEGMENTATION PROTEIN PAIRED
8849	21816	35236	0.57	8.5E-01	AJ243213.1	NT	Homo sapiens partial 5-HT4 receptor gene, exons 2 to 5
10716	23637	37129	1.35	8.5E-01	AB008789.1	NT	Cyertidium caldarium gene for SigC, complete cds
10715	23637	37130	1.35	8.5E-01	AB008789.1	NT	Cyertidium caldarium gene for SigC, complete cds
12565	25859		2.25	8.5E-01	11418543	NT	Homo sapiens human immunodeficiency virus type 1 enhancer-binding protein 1 (HIVBP1), mRNA
12572	25283		1.37	8.5E-01	9507008	NT	Rattus norvegicus protein tyrosine phosphatase, non-receptor type 5 (Ptpn5), mRNA
4782	17802	30683	0.85	8.4E-01	AF083975.2	NT	Fowl adenovirus 8, complete genome
5571	25841	31827	2.68	8.4E-01	L78726.1	NT	Human fibroblast growth factor receptor 3 (FGFR3) gene, intron 7
5571	25841	31828	2.68	8.4E-01	L78726.1	NT	Human fibroblast growth factor receptor 3 (FGFR3) gene, intron 7
8107	21044	34449	0.61	8.4E-01	AF051142.1	NT	Mamestra brassicae phenolase binding protein 2 precursor (PBP2) mRNA, complete cds
10317	23241		3.25	8.4E-01	AJ248287.1	NT	Pyrococcus abyssi complete genome, segment 5/6
12005	24882	38478	1.54	8.4E-01	M55584.1	NT	Human collagenase type IV (CLG4) gene, exon 4
743	13804	28743	3.01	8.3E-01	M83437.1	NT	Thermus thermophilus cytochrome c-552 (cycA) and CycB (cycB) genes, complete cds
3111	16168	28078	3.15	8.3E-01	AL161608.2	NT	Arabidopsis thaliana DNA chromosome 4, contig fragment No. 18
3823	16863	28078	0.83	8.3E-01	AB010878.1	NT	Nicotiana tabacum mRNA for chloroplast ribosomal protein L10, complete cds
4040	17078	28978	3.15	8.3E-01	Y19177.1	NT	Streptomyces antibioticus polyketide biosynthetic gene cluster
5341	18446	31189	2.42	8.3E-01	AL161540.2	NT	Arabidopsis thaliana DNA chromosome 4, contig fragment No. 40

Page 35 of 546
Table 4

Single Exon Probes Expressed in Bone Marrow

Probe SEQ ID NO:	Exon SEQ ID NO:	ORF SEQ ID NO:	Expression Signal	Most Similar (Top) Hit BLAST E Value	Top Hit Accession No.	Top Hit Database Source	Top Hit Descriptor
10028	22855		4.53	8.3E-01	AI791952.1	EST_HUMAN	nr0112.y6 NCI_CGAP_C69 Homo sapiens cDNA clone IMAGE:1076485 5' similar to contains THR.H THR repetitive element;
10469	23391	36886	1.1	8.3E-01	AF068070.1	NT	Drosophila melanogaster Lsr1 homolog mRNA, complete cds
10579	23501	36903	3.97	8.3E-01	AF108133.1	NT	Mus musculus neuro-d4 gene, exons 3 through 12 and partial cds
11033	23997	37524	2.97	8.3E-01	AE000903.1	NT	Methanobacterium thermotrophicum from bases 1270510 to 1283409 (section 109 of 148) of the complete genome
11050	24013		1.92	8.3E-01	7212472	NT	Phytophthora infestans mitochondrion, complete genome
11634	24571	38135	2.22	8.3E-01	AF020503.1	NT	Homo sapiens FRA3B common fragile region, diadenosine triphosphate hydrolase (FHIT) gene, exon 5
2068	15085	28103	2.24	8.2E-01	AB000489.1	NT	Rattus norvegicus mRNA for RPHO-1, complete cds
2101	15118		1.08	8.2E-01	AF145589.1	NT	Mus musculus trophinin (Tnn) gene, complete cds
2688	15684		1.06	8.2E-01	AW376980.1	EST_HUMAN	IL3-CT0219-161199-031-Q08 CT0219 Homo sapiens cDNA
3918	16958	29871	0.75	8.2E-01	AB014574.1	NT	Homo sapiens mRNA for KIAA0674 protein, partial cds
3940	16980	29895	0.92	8.2E-01	AF063417.1	NT	Tanystylum orbiculare elongation factor 1-alpha mRNA, partial cds
6800	19854	33139	0.57	8.2E-01	X95283.1	NT	G.gallus mRNA for C-Serrat-1 protein
6800	19854	33140	0.57	8.2E-01	X95283.1	NT	G.gallus mRNA for C-Serrat-1 protein
6941	20165	33488	0.85	8.2E-01	AJ010142.1	NT	Ananitis muscaria mRNA for SCIL25 protein
7082	20103	33414	3.69	8.2E-01	AW378433.1	EST_HUMAN	CM4-HT0243-081199-037-e01 HT0243 Homo sapiens cDNA
7484	25682	33807	4.38	8.2E-01	Z12128.1	NT	S. cerevisiae MET, LEU4, and POL1 genes encoding MET4 protein, alpha-isopropylmalate (alpha-IPM) synthetase (partial), and DNA polymerase alpha (partial)
8787	21754	35176	0.58	8.2E-01	BE263145.1	EST_HUMAN	60114486F2 NIH_MGC_19 Homo sapiens cDNA clone IMAGE:3160412 5'
10385	23307	36785	0.68	8.2E-01	AB014530.1	NT	Homo sapiens mRNA for KIAA0630 protein, partial cds
10419	23341	36827	1.59	8.2E-01	AF052659.1	NT	Homo sapiens thiredoxin-related protein mRNA, complete cds
10583	23505	36908	0.59	8.2E-01	AF223888.1	NT	Oncorhynchus tshawytscha isolate T-20 somatolactin precursor gene, exon 1
10583	23505	36909	0.59	8.2E-01	AF223888.1	NT	Oncorhynchus tshawytscha isolate T-20 somatolactin precursor gene, exon 1
10761	23673	37169	3.84	8.2E-01	Q9J170	SWISSPROT	MCKUSICK-KAUFMAN/BARDET-BIEDL SYNDROMES PUTATIVE CHAPERONIN
10751	23673	37170	3.84	8.2E-01	Q9J170	SWISSPROT	MCKUSICK-KAUFMAN/BARDET-BIEDL SYNDROMES PUTATIVE CHAPERONIN
11657	24835	38432	2.97	8.2E-01	L10127.1	NT	Mollusca conglusum virus type 1 ORF1 and ORF2 DNA
12038	24913	38507	4.82	8.2E-01	P10383	SWISSPROT	OVARIAN TUMOR LOCUS PROTEIN
12043	24918	38514	6.1	8.2E-01	H87398.1	EST_HUMAN	yw14402.r1 Soares_placenta_8hrsweeks_2NltHP8tc9W Homo sapiens cDNA clone IMAGE:252185 5' similar to gb:M36072.80S RIBOSOMAL PROTEIN L7A (HUMAN);
12586	25290	31781	2.98	8.2E-01	AJ001261.1	NT	Mus musculus mRNA for NIP-SNAP2 protein
2769	15761		1.48	8.1E-01	AF191839.1	NT	Mus musculus TANK binding kinase TBK1 (Tbk1) mRNA, complete cds
3472	16518	29439	3.67	8.1E-01	AF055068.1	NT	Homo sapiens MHC class 1 region

Page 36 of 546
Table 4

Single Exon Probes Expressed in Bone Marrow

Probe SEQ ID NO:	Exon SEQ ID NO:	ORF SEQ ID NO:	Expression Signal	Most Similar (Top) Hit BLAST E Value	Top Hit Accession No.	Top Hit Database Source	Top Hit Descriptor
3472	18518	28440	3.67	8.1E-01	AF055068.1	NT	Homo sapiens MHC class 1 region
5792	18984	32068	0.51	8.1E-01	Q01727	SWISSPROT	MELANOCYTE STIMULATING HORMONE RECEPTOR (MSH-R) (MELANOTROPIN RECEPTOR)
6448	19513	32763	0.84	8.1E-01	U16780.1	NT	(MELANOCORTIN-1 RECEPTOR) (MC1-R)
6777	19832	33114	2.47	8.1E-01	Q13491	SWISSPROT	Mus musculus putative collagen alpha-2 (X1) chain (COL11A2) gene, partial cds
6777	19832	33115	2.47	8.1E-01	Q13491	SWISSPROT	NEURONAL MEMBRANE GLYCOPROTEIN M6-B
7755	20708	34077	0.55	8.1E-01	Q47477	SWISSPROT	NEURONAL MEMBRANE GLYCOPROTEIN M6-B
							CYTCHROME B
8243	21212	34618	1.12	8.1E-01	AF02713.2	NT	Drosophila melanogaster putative inorganic phosphatase cotransporter (Plocat) gene, partial cds; putative sodium channel (Nach) and putative amylase-related protein (Amyrel) genes, complete cds; and putative serine-enriched protein (gprs) gene, partial cd
8243	21212	34619	1.12	8.1E-01	AF02713.2	NT	Drosophila melanogaster putative inorganic phosphatase cotransporter (Plocat) gene, partial cds; putative sodium channel (Nach) and putative amylase-related protein (Amyrel) genes, complete cds; and putative serine-enriched protein (gprs) gene, partial cd
8956	21922	35349	0.91	8.1E-01	AP001517.1	NT	Bacillus halodurans genomic DNA, section 11/14
8956	21922	35350	0.91	8.1E-01	AP001517.1	NT	Bacillus halodurans genomic DNA, section 11/14
							xx01103.x1 NCL_CGAP_Kid11 Homo sapiens cDNA clone IMAGE2692469 3' similar to SW-1.YAR_MOUSE Q08288 CELL GROWTH REGULATING NUCLEAR PROTEIN. ;contains MER22.b1 PTR5 repetitive element ;
9122	22088	35516	1.13	8.1E-01	AW242847.1	EST_HUMAN	PROBABLE E4 PROTEIN
10484	23406	36902	0.64	8.1E-01	P08425	SWISSPROT	KK8872F Human fetal heart, Lambda ZAP Express Homo sapiens cDNA clone KK8872 5' similar to EST(CLONE C-0PE11)
10776	23697	37185	0.42	8.1E-01	N84541.1	EST_HUMAN	RCO-TN0080-220800-025-d10 TN0080 Homo sapiens cDNA
11812	24697	38277	4.05	8.1E-01	BE938593.1	EST_HUMAN	RCO-TN0080-220800-025-d10 TN0080 Homo sapiens cDNA
11812	24697	38278	4.05	8.1E-01	BE938593.1	EST_HUMAN	RCO-TN0080-220800-025-d10 TN0080 Homo sapiens cDNA
12298	25109	31839	1.73	8.1E-01	AE000171.1	NT	Thermotoga maritima section 23 of 136 of the complete genome
178	13278		3.32	8.0E-01	AJ271510.1	NT	Staphylococcus aureus partial pta gene for phosphatase acetyltransferase allele 15
288	13383	26310	5.97	8.0E-01	AJ132772.1	NT	Bos taurus tubb and rtf genes
2051	15070		1.72	8.0E-01	BF30962.1	EST_HUMAN	602072473F1 NCL_CGAP_Bim67 Homo sapiens cDNA clone IMAGE:4215091 5'
3093	16151	29065	1.41	8.0E-01	AF127897.1	NT	Salmonella enteritidis diffractory receptor (SBOZ7) gene, partial cds
3324	16375	28298	1.3	8.0E-01	AB006183.1	NT	Mus musculus gene for ovalbumin glycoprotein, complete cds
3717	16760		2.36	8.0E-01	AL162758.2	NT	Neisseria meningitidis serogroup A strain Z2491 complete genome; segment 777
4563	17586	30478	6.45	8.0E-01	X83739.2	NT	G.gallus mRNA for nicotinic acetylcholine receptor (nAChR) beta 3 subunit
8322	21291		2.31	8.0E-01	AW901489.1	EST_HUMAN	RCO-NN1012-270300-021-h08 NN1012 Homo sapiens cDNA
8870	21837	35259	1.05	8.0E-01	Y11095.1	NT	Rice stripe virus RNA 3
11303	24253	37778	1.58	8.0E-01	Q82783	SWISSPROT	CREB-BINDING PROTEIN

Page 37 of 546
Table 4

Single Exon Probes Expressed in Bone Marrow

Probe SEQ ID NO:	Exon SEQ ID NO:	ORF SEQ ID NO:	Expression Signal	Most Similar (Top) Hit BLAST E Value	Top Hit Accession No.	Top Hit Database Source	Top Hit Descriptor
454	13527	28457	0.78	7.9E-01	D11478.1	NT	Lymantia dispar nuclear polyhedrosis virus gene for DNA polymerase, complete cds
716	13778		0.78	7.9E-01	AE002130.1	NT	Ureaplasma urealyticum section 31 of 69 of the complete genome
1609	14641		23.05	7.9E-01	AB040885.1	NT	Homo sapiens mRNA for KIAA1452 protein, partial cds
1663	14695		1.28	7.9E-01	U32739.1	NT	Haemophilus influenzae Rd section 54 of 163 of the complete genome
2272	15285	28311	7.38	7.9E-01	AB004816.1	NT	Onyctolagus cuniculus mRNA for mitsugumin29, complete cds
2273	15286	28312	2.36	7.9E-01	AF130459.1	NT	Danio rerio Tnp4-associated protein Tap1A (tap1A) mRNA, complete cds
3528	16574	29497	3	7.9E-01	AF228664.1	NT	Gallus gallus SOX8 transcription factor (SOX8) mRNA, complete cds
4330	17358		0.88	7.9E-01	BE263612.1	EST_HUMAN	601182033F1 NIH_MGC. 7 Homo sapiens cDNA clone IMAGE:3535785 5'
4642	17663	30551	1.15	7.9E-01	6763745	NT	Mus musculus embigin (Emb), mRNA
4642	17663	30552	1.15	7.9E-01	6763745	NT	Mus musculus embigin (Emb), mRNA
5244	18252	31123	1.03	7.9E-01	AF228843.1	NT	Mus musculus WNT-2 gene, partial cds; putative ankyrin-related protein and cystic fibrosis transmembrane conductance regulator (CFTR) genes, section 1 of 2 of the complete cds; and unknown gene
6479	19544	32792	0.76	7.9E-01	D38145.1	NT	Human mRNA for prostacyclin synthase, complete cds
8448	21415	34828	3.37	7.9E-01	X00998.1	NT	P. sativum GR gene
8905	22857	36319	4.3	7.9E-01	U01912.1	NT	Giardia lamblia variant-specific surface protein G3M-B (vspG3M-B) mRNA, partial cds
10409	23331	36816	4.3	7.9E-01	P19719	SWISSPROT	SMALL HYDROPHOBIC PROTEIN
10451	23373	36864	0.82	7.9E-01	AV700860.1	EST_HUMAN	AV700860 GKC Homo sapiens cDNA clone GKCDRE12 3'
10877	23787	37298	0.82	7.9E-01	AB000631.1	NT	Streptococcus mutans DNA for sigma 42 protein, dTDP-4-keto-L-rhamnose reductase, complete cds
11350	24300		2.81	7.9E-01	7682471	NT	Homo sapiens KIAA1072 protein (KIAA1072), mRNA
11546	24487	38041	2.22	7.9E-01	P19022	SWISSPROT	NEURAL-CADHERIN PRECURSOR (N-CADHERIN)
877	13632		1.98	7.9E-01	Z43785.1	EST_HUMAN	HSC1KH041 normalized infant brain cDNA Homo sapiens cDNA clone o-1kh04
2283	15298	28320	6.12	7.9E-01	AW959567.1	EST_HUMAN	EST371637 MAGE resequences, MAGF Homo sapiens cDNA
4730	17750	30842	1.33	7.9E-01	U87905.1	NT	Rattus norvegicus transmembrane receptor Unc5H1 mRNA, complete cds
5076	18085		0.8	7.9E-01	AW753363.1	EST_HUMAN	RC3-CT0254-130100-023-c02 CT0254 Homo sapiens cDNA
6187	19282	32498	2.5	7.9E-01	AF115856.1	NT	Sphenodon punctatus alpha endase mRNA, partial cds
6344	19413	32855	1.04	7.9E-01	P06231	SWISSPROT	INTERLEUKIN-6 PRECURSOR (IL-6) (B-CELL STIMULATORY FACTOR 2) (BSF-2) (INTERFERON BETA-2) (HYBRIDOMA GROWTH FACTOR)
6006	19664	32939	0.72	7.9E-01	AL445068.1	NT	Thermoplasma acidophilum complete genome; segment 4/5
8836	21803	35221	1.29	7.9E-01	BF108927.1	EST_HUMAN	754d05.x1 Soares_NSJ_F8_9W_OT_PA_P_S1 Homo sapiens cDNA clone IMAGE:3525176 3'
9589	22551	36002	1.34	7.9E-01	Y10169.1	NT	D discoideum racGAP gene
9687	22640	36097	0.52	7.9E-01	4826873	NT	Homo sapiens nucleoporin 214kD (CIN) (NUP214), mRNA
10483	23405		1.01	7.9E-01	Q25452	SWISSPROT	MUSCLE CALCIUM CHANNEL ALPHA-1 SUBUNIT (MDL-ALPHA1)

Page 38 of 546
Table 4

Single Exon Probes Expressed In Bone Marrow

Probe SEQ ID NO:	Exon SEQ ID NO:	ORF SEQ ID NO:	Expression Signal	Most Similar (Top) Hit BLAST E Value	Top Hit Accession No.	Top Hit Database Source	Top Hit Descriptor
12559	25848		2.32	7.6E-01	28280.1	NT	Arabidopsis thaliana 1-aminio-1-cyclopropanecarboxylate synthase (ACS5) gene, complete cds
145	13248	28177	6.69	7.7E-01	AF184345.1	NT	Lycopersicon hirsutum ADP-glucose pyrophosphorylase large subunit (AGP-L1) mRNA, complete cds
727	13788		3.26	7.7E-01	AF050157.1	NT	Mus musculus major histocompatibility locus class II region: major histocompatibility protein class II alpha chain (Aa1pha) and major histocompatibility protein class II beta chain (Ib2eta) genes, complete cds; butyrophilin-like (NG9), butyrophilin-IP
2724	15718	28736	2	7.7E-01	O33915	SWISSPROT	CITRATE SYNTHASE
3368	16418		0.76	7.7E-01	8393408	NT	Homo sapiens UDP-N-acetyl-alpha-D-galactosamine:polypeptide N-acetylgalactosaminyltransferase 7 (GalNAC-T7) (GALNAC-T7), mRNA
3614	16958	29576	3.83	7.7E-01	AF118085.1	NT	Homo sapiens PRO1975 mRNA, complete cds
4425	17452	30343	3.04	7.7E-01	AF189488.1	NT	Coturnix coturnix japonica sub-species japonica beta-actin mRNA, partial cds
4425	17452	30344	3.04	7.7E-01	AF189488.1	NT	Coturnix coturnix japonica sub-species japonica beta-actin mRNA, partial cds
5640	18736	31889	1.25	7.7E-01	P16553	SWISSPROT	RAFFINOSE INVERTASE (INVERTASE)
5640	18736	31900	1.25	7.7E-01	P16553	SWISSPROT	RAFFINOSE INVERTASE (INVERTASE)
6066	19147	32359	0.53	7.7E-01	R08600.1	EST_HUMAN	yf24b02.s1 Soares fetal liver spleen 1NFLS Homo sapiens cDNA clone IMAGE:127755 3'
10204	23129	36816	0.72	7.7E-01	AB021134.1	NT	Dephnia magna hemoglobin gene cluster (dhb3, dhb1 and dhb2 genes), complete cds
12447	25207		5.53	7.7E-01	11497621	NT	Archaeoglobus fulgidus, complete genome
8218	19292	32525	4.49	7.6E-01	AF059510.1	NT	Arabidopsis thaliana 3-methylcrotonyl-CoA carboxylase non-biotinylated subunit (MCCB) mRNA, complete cds
8218	19292	32526	4.49	7.6E-01	AF059510.1	NT	Arabidopsis thaliana 3-methylcrotonyl-CoA carboxylase non-biotinylated subunit (MCCB) mRNA, complete cds
6867	19724	32899	0.65	7.6E-01	P37838	SWISSPROT	MATING-TYPE PROTEIN A-ALPHA Z4
7029	18361	31248	0.88	7.6E-01	A1253399.1	EST_HUMAN	eq14b12.x1 Stanley Frontal NS pool 2 Homo sapiens cDNA clone IMAGE:2030879
7029	18361	31282	0.88	7.6E-01	A1253399.1	EST_HUMAN	eq14b12.x1 Stanley Frontal NS pool 2 Homo sapiens cDNA clone IMAGE:2030879
7252	19987	33285	0.89	7.6E-01	U72487.1	NT	Rattus norvegicus calcium-independent alpha-lactalbumin receptor mRNA, complete cds
8400	21369	34778	1.33	7.6E-01	AF146783.2	NT	Mus musculus neuromedin U precursor (Nmu) gene, partial cds; tPhLP (Tphlp) gene, partial cds; CLOCK (Clock) gene, complete cds; PFT27 (Pft27) gene, complete cds; and H5AR (H5ar) gene, complete cds
8464	21433	34849	1.92	7.6E-01	6857752	NT	Mus musculus adillin (Adill-pending), mRNA
8464	21433	34850	1.92	7.6E-01	6857752	NT	Mus musculus adillin (Adill-pending), mRNA
8668	21636	35058	0.43	7.6E-01	Q01098	SWISSPROT	GLUTAMATE [NMDA] RECEPTOR SUBUNIT EPSILON 3 PRECURSOR (N-METHYL D-ASPARTATE RECEPTOR SUBTYPE 2C) (NR2C) (NMDAR2C)
8668	21636	35059	0.43	7.6E-01	Q01098	SWISSPROT	GLUTAMATE [NMDA] RECEPTOR SUBUNIT EPSILON 3 PRECURSOR (N-METHYL D-ASPARTATE RECEPTOR SUBTYPE 2C) (NR2C) (NMDAR2C)

Table 4

Single Exon Probes Expressed In Bone Marrow

Probe SEQ ID NO:	Exon SEQ ID NO:	ORF SEQ ID NO:	Expression Signal	Most Similar (Top) Hit BLAST E Value	Top Hit Accession No.	Top Hit Database Source	Top Hit Descriptor
9318	22283	35714	1.01	7.6E-01	6753577	NT	Mus musculus cytochrome P450, 2B9, phenobarbital inducible, type a (Cyp2b9), mRNA
9334	22578	36028	5.25	7.6E-01	P30372	SWISSPROT	MUSCARINIC ACETYLCHOLINE RECEPTOR M2
9334	22578	36028	5.25	7.6E-01	P30372	SWISSPROT	MUSCARINIC ACETYLCHOLINE RECEPTOR M2
11685	24651	38229	2.09	7.6E-01	X86347.1	NT	H. aspersa mRNA for neurofilament NF70
11685	24651	38230	2.09	7.6E-01	X86347.1	NT	H. aspersa mRNA for neurofilament NF70
12020	24697		3.05	7.6E-01	AL161592.2	NT	Arabidopsis thaliana DNA chromosome 4, contig fragment No. 88
12203	25045		3.8	7.6E-01	AB020702.1	NT	Homo sapiens mRNA for KIAA0886 protein, partial cds
514	13585		1.67	7.5E-01	AL163301.2	NT	Homo sapiens chromosome 21 segment HS21C101
588	13654	28568	1.01	7.5E-01	AF020503.1	NT	Homo sapiens FRA3B common fragile region, diadenosine triphosphate hydrolase (FHIT) gene, exon 5
7764	20717	34080	0.78	7.5E-01	AF052730.1	NT	Drosophila melanogaster tyrosine kinase receptor protein (eph) mRNA, complete cds
12616	25247		4.93	7.5E-01	AF163151.2	NT	Homo sapiens dentin sialophosphoprotein precursor (DSPP) gene, complete cds
13027	25575	31697	1.57	7.5E-01	AE000823.1	NT	Methanobacterium thermoautotrophicum from bases 31750 to 328792 (section 29 of 148) of the complete genome
1132	14175	27124	1.35	7.4E-01	AI698148.1	EST_HUMAN	h114b09.x1 NCI_CGAP_Bm25 Homo sapiens cDNA clone IMAGE:2167577 3' similar to contains Alu repetitive element; contains element MIR repetitive element;
2350	15359	28381	0.93	7.4E-01	AB011108.1	NT	Homo sapiens mRNA for KIAA0634 protein, partial cds
3739	16781	29883	18.81	7.4E-01	AF112538.1	NT	Malva pusilla actin (Act1) mRNA, complete cds
3919	16959	29872	2.01	7.4E-01	AF193310.1	NT	Vibrio cholerae phage CTXphi Calcutta-rsR-a (rsR-a) and Calcutta-rsR-b (rsR-b) genes, complete cds
4340	17367	30260	8.99	7.4E-01	AL163248.2	NT	Homo sapiens chromosome 21 segment HS21C046
8176	21146	34563	1.04	7.4E-01	AL161551.2	NT	Arabidopsis thaliana DNA chromosome 4, contig fragment No. 51
8176	21146	34554	1.04	7.4E-01	AL161551.2	NT	Arabidopsis thaliana DNA chromosome 4, contig fragment No. 51
8880	21946	36370	0.78	7.4E-01	BF346286.1	EST_HUMAN	602018456F1 NCI_CGAP_Bm67 Homo sapiens cDNA clone IMAGE:4154340 5'
9063	22029		0.64	7.4E-01	U87980.1	NT	Rattus norvegicus leukocyte common antigen receptor (LAR) gene, trans-spliced alternative untranslated exon
9452	22416	36854	7.09	7.4E-01	BE747503.1	EST_HUMAN	601573028F1 NIH_MGC_9 Homo sapiens cDNA clone IMAGE:3834174 5'
9511	22474	36918	1.14	7.4E-01	AA187988.1	EST_HUMAN	zp67h01.a1 Strabane endothelial cell 837223 Homo sapiens cDNA clone IMAGE:025297 3' similar to SW:TCPO_MOUSE_P42932 T-COMPLEX PROTEIN 1, THETA SUBUNIT;
10767	23698	37185	0.74	7.4E-01	11424933	NT	Homo sapiens NY-REN-45 antigen (LOC51133), mRNA
12171	25021		4.46	7.4E-01	6753217	NT	Mus musculus complement component 1 inhibitor (C1inh), mRNA
12284	25101		1.64	7.4E-01	AI472841.1	EST_HUMAN	h13h01.x1 NCI_CGAP_Lyn5 Homo sapiens cDNA clone IMAGE:2043985 3'
3999	17038		0.84	7.3E-01	AF000082.1	NT	Aeropyrum pernix genome DNA, section 57
4647	17668	30555	0.97	7.3E-01	AE001166.1	NT	Borrelia burgdorferi (section 52 of 70) of the complete genome

Page 40 of 546
Table 4

Single Exon Probes Expressed in Bone Marrow

Probe SEQ ID NO:	Exon SEQ ID NO:	ORF SEQ ID NO:	Expression Signal	Most Similar (Top) Hit BLAST E Value	Top Hit Accession No.	Top Hit Database Source	Top Hit Descriptor
4729	17749	30641	4.57	7.3E-01	AF225421.1	NT	Homo sapiens HT017 mRNA, complete cds
5135	18144	31024	1.18	7.3E-01	O43103	SWISSPROT	FERRIC-IRON SIDEROPORE PEPTIDE SYNTHETASE
6761	18815	33094	6.03	7.3E-01	L35772.1	NT	Mus musculus antigen (CD72) gene
6761	18816	33095	6.03	7.3E-01	L35772.1	NT	Mus musculus antigen (CD72) gene
7301	25679	33608	0.82	7.3E-01	AJ01418.1	NT	Lycopodium esculentum mRNA for ubiquitin activating enzyme
7682	20650	34014	0.53	7.3E-01	Z14133.1	NT	D.melanogaster Cho mRNA for clathrin heavy chain
7794	20746	34119	7.46	7.3E-01	M26511.1	NT	V.alginolyticus sucrose (scrB) gene, complete cds
7794	20746	34120	7.46	7.3E-01	M26511.1	NT	V.alginolyticus sucrose (scrB) gene, complete cds
8115	21052	34450	0.51	7.3E-01	U34631.1	NT	Mus musculus alpha-4 integrin gene, exon 7
11759	24687	38267	3.11	7.3E-01	AA678019.1	EST_HUMAN	z25b08.s1 Soares fetal liver spleen_1N1FLS S1 Homo sapiens cDNA clone IMAGE:431788 3'
11759	24687	38268	3.11	7.3E-01	AA678019.1	EST_HUMAN	z25b08.s1 Soares fetal liver spleen_1N1FLS S1 Homo sapiens cDNA clone IMAGE:431788 3'
832	19889		2.03	7.2E-01	L28281.1	NT	Rattus norvegicus initiation factor-2 kinase (eIF-2a) mRNA, complete cds
1974	14895	27997	3.23	7.2E-01	X79140.1	NT	N.tabacum Neif-4A13 mRNA
2468	15472	28465	1.91	7.2E-01	AB009805.1	NT	Gallus gallus gene for melanocortin 2-receptor, complete cds
3080	16137	29048	1.47	7.2E-01	AF198100.1	NT	Fowlpox virus, complete genome
3467	16513	29434	2.44	7.2E-01	AF065508.1	NT	Giardia intestinalis variant-specific surface protein (vsp417-6) gene, vsp417-6/A-1 allele, complete cds
4803	17820	30714	2.69	7.2E-01	D90314.1	NT	L.messerioides gene for sucrose phosphorylase (EC 2.4.1.7)
5159	18168	31046	1.57	7.2E-01	AF198779.1	NT	Homo sapiens transcription factor IGHM enhancer 3, JIM1 protein, JIM4 protein, JIM5 protein, T54 protein, JIM10 protein, A4 differentiation-dependent protein, triple LIM domain protein 6, and synaptophysin genes, complete cds; and L-type calcium channel α
5159	18168	31047	1.57	7.2E-01	AF198779.1	NT	Homo sapiens transcription factor IGHM enhancer 3, JIM1 protein, JIM4 protein, JIM5 protein, T54 protein, JIM10 protein, A4 differentiation-dependent protein, triple LIM domain protein 6, and synaptophysin genes, complete cds; and L-type calcium channel α
7421	20388	33739	0.78	7.2E-01	U68833.1	NT	Solanum tuberosum cold-stress inducible protein (C17) gene, complete cds
8787	21764	35186	1.24	7.2E-01	AF236061.1	NT	Oryctolagus cuniculus RING-finger binding protein mRNA, partial cds
9814	22279	37123	0.52	7.2E-01	AF743773.1	EST_HUMAN	AV743773 CB Homo sapiens cDNA clone CBMAFD06 5'
10705	23627	37123	2.59	7.2E-01	BF670061.1	EST_HUMAN	602118381F1 NIH_MGC_56 Homo sapiens cDNA clone IMAGE:4276381 5'
11081	24051	37574	3.38	7.2E-01	U82623.1	NT	Rattus norvegicus cytochrome mRNA, complete cds
12523	18343	31291	1.43	7.2E-01	U02568.1	NT	Drosophila melanogaster polyprotein antigen precursor (Dva) mRNA, complete cds
12700	25360		5.56	7.2E-01	AP000063.1	NT	Aeropyrum pernix genomic DNA, section 6/7
683	13755	26885	13.3	7.1E-01	D21070.1	NT	Rana catesbeiana mRNA for bullfrog skeletal muscle calcium release channel (ryanodine receptor) alpha isoform(RyR1), complete cds
3075	16132	28045	13.21	7.1E-01	AJ270777.1	NT	Homo sapiens partial TCF-4 gene for T-cell transcription factor-4, exons 15-16

Page 41 of 546
Table 4
Single Exon Probes Expressed in Bone Marrow

Probe SEQ ID NO:	Exon SEQ ID NO:	ORF SEQ ID NO:	Expression Signal	Most Similar (Top) BLAST E Value	Top Hit Accession No.	Top Hit Database Source	Top Hit Descriptor
4236	17265	30151	3.29	7.1E-01	7905360	NT	Mus musculus otogelin (Otog), mRNA
4236	17265	30152	3.29	7.1E-01	7905360	NT	Mus musculus otogelin (Otog), mRNA
6058	19139	32350	1.55	7.1E-01	BF681034.1	EST_HUMAN	602155438F1 NIH_MGC 83 Homo sapiens cDNA clone IMAGE:4266344 5'
6058	19139	32351	1.55	7.1E-01	BF681034.1	EST_HUMAN	602155438F1 NIH_MGC 83 Homo sapiens cDNA clone IMAGE:4266344 5'
7137	20113	33428	8.92	7.1E-01	U36232.1	NT	Drosophila melanogaster 6-pyruvyltetrahydropterin synthase (pr) gene, complete cds
8533	21501	34918	0.48	7.1E-01	H54244.1	EST_HUMAN	y989d09.s1 Soares fetal liver spleen 1N1FLS Homo sapiens cDNA clone IMAGE:202861 3'
9088	22054	35477	0.85	7.1E-01	BE074185.1	EST_HUMAN	RC1-BT0567-301289-011-d09 BT0567 Homo sapiens cDNA
9088	22054	35478	0.85	7.1E-01	BE074185.1	EST_HUMAN	RC1-BT0567-301289-011-d09 BT0567 Homo sapiens cDNA
10214	23189	36828	1.28	7.1E-01	BE904405.1	EST_HUMAN	601496330F1 NIH_MGC_70 Homo sapiens cDNA clone IMAGE:3898495 5'
10774	23695	37193	1.13	7.1E-01	M12891.1	NT	Human T-cell receptor gamma-chain J2 gene
12499	25773		2.34	7.1E-01	AA421492.1	EST_HUMAN	zu08n11.s1 Soares testis_NHT Homo sapiens cDNA clone IMAGE:731109 3'
1233	14270	27228	1.13	7.0E-01	AB014514.1	NT	Homo sapiens mRNA for KIAA0814 protein, partial cds
1233	14270	27230	1.13	7.0E-01	AB014514.1	NT	Homo sapiens mRNA for KIAA0814 protein, partial cds
2455	15460	28482	1.09	7.0E-01	N62412.1	EST_HUMAN	y273e07.s1 Soares multiple_sclerosis_2NBH-MSP Homo sapiens cDNA clone IMAGE:286708 3' similar to contains Alu repetitive element;
2455	15460	28483	1.09	7.0E-01	N62412.1	EST_HUMAN	y273e07.s1 Soares multiple_sclerosis_2NBH-MSP Homo sapiens cDNA clone IMAGE:286708 3' similar to contains Alu repetitive element;
5098	18108		2.11	7.0E-01	AL163301.2	NT	Homo sapiens chromosome 21 segment HS21C101
6082	19143		0.95	7.0E-01	AB021318.1	NT	Arabidopsis thaliana mRNA for chlorophyll b synthase, complete cds
8721	21889		8.51	7.0E-01	AE000253.1	NT	Escherichia coli K-12 MG1655 section 143 of 400 of the complete genome
9671	22624	36077	0.52	7.0E-01	U53898.1	NT	Clostridium acetobutylicum mannitol-specific phosphotransferase system (PTS) system, mtdA, mtdR, mtdF, and mtdD genes, complete cds
9671	22624	36078	0.52	7.0E-01	U53898.1	NT	Clostridium acetobutylicum mannitol-specific phosphotransferase system (PTS) system, mtdA, mtdR, mtdF, and mtdD genes, complete cds
11454	24397	37843	1.71	7.0E-01	AV763842.1	EST_HUMAN	AV763842 MDS Homo sapiens cDNA clone MDSCH04 5'
11454	24397	37844	1.71	7.0E-01	AV763842.1	EST_HUMAN	AV763842 MDS Homo sapiens cDNA clone MDSCH04 5'
13035	26811	31528	1.35	7.0E-01	9630464	NT	Bacteriophage N15 v10n, complete genome
971	14023	26976	12.59	8.9E-01	U69674.1	NT	Candida albicans squalene epoxidase (CAERG1) gene, complete cds and translational regulator gene, partial cds
971	14023	26977	12.59	8.9E-01	U69674.1	NT	Candida albicans squalene epoxidase (CAERG1) gene, complete cds and translational regulator gene, partial cds
1313	14349	27315	2.22	8.9E-01	AA593530.1	EST_HUMAN	rn28a09.s1 NCI_CGAP_Gas1 Homo sapiens cDNA clone IMAGE:1085176 3'
3233	16288	28210	1.8	8.9E-01	AE002271.2	NT	Chlamydia muridarum, section 3 of 85 of the complete genome
5249	18257	31126	0.92	8.9E-01	AV714502.1	EST_HUMAN	AV714502 DCB Homo sapiens cDNA clone DCBATD12 5'

Page 42 of 546
Table 4

Single Exon Probes Expressed in Bone Marrow

Probe SEQ ID NO:	Exon SEQ ID NO:	ORF SEQ ID NO:	Expression Signal	Most Similar (Top) Hit BLAST E Value	Top Hit Accession No.	Top Hit Database Source	Top Hit Descriptor
5878	18867	32158	0.8	6.9E-01	AB035682.1	NT	Branchiostoma belcheri BbNA3 mRNA for notochord actin, complete cds
6104	19183	32402	0.55	6.9E-01	Y18278.1	NT	Drosophila melanogaster mRNA for A-kinase anchor protein DAKAP550, partial
6506	19570	32822	1.5	6.9E-01	BE298188.1	EST_HUMAN	60117733F1 NIH_MGC_17 Homo sapiens cDNA clone IMAGE:3532328 5'
8312	21281	34682	3.39	6.9E-01	AL161573.2	NT	Arabidopsis thaliana DNA chromosome 4, contig fragment No. 69
8312	21281	34683	3.39	6.9E-01	AL161573.2	NT	Arabidopsis thaliana DNA chromosome 4, contig fragment No. 69
9526	22489		0.73	6.9E-01	AF116046.1	NT	Entamoeba dispar cation transporting ATPase (atpase) gene, partial cds
10053	22980	38447	0.55	6.9E-01	AF206319.1	NT	Musa acuminata pectate lyase 1 (PL1) mRNA, complete cds
10053	22980	38448	0.55	6.9E-01	AF206319.1	NT	Musa acuminata pectate lyase 1 (PL1) mRNA, complete cds
11587	24525	38081	2.03	6.9E-01	D89013.1	NT	Homo sapiens DAN gene, complete cds
11587	24525	38082	2.03	6.9E-01	D89013.1	NT	Homo sapiens DAN gene, complete cds
12145	25768		3.91	6.9E-01	Q99958	SWISSPROT	FORKHEAD BOX PROTEIN C2 (FORKHEAD-RELATED PROTEIN FKHL14) (MESENCHYME FORK HEAD PROTEIN 1) (MFL-1 PROTEIN) (TRANSCRIPTION FACTOR FKHL14)
958	14011	26964	1.52	6.8E-01	AF017784.1	NT	Glardia intestinalis carbamate kinase gene, complete cds
2682	15978		1.26	6.8E-01	D90917.1	NT	Synechocystis sp. PCC6803 complete genome, 2727, 3418852-3573470
2840	14951	27627	1.5	6.8E-01	AA854475.1	EST_HUMAN	aj75a05.s1 Soares_parathyroid_tumor_NbHPA Homo sapiens cDNA clone IMAGE:1402256 3' similar to gb:X568411_mn1 ALCOHOL DEHYDROGENASE CLASS II P1 CHAIN (HUMAN);
4602	17623	30516	1.76	6.8E-01	J00762.1	NT	Rat(hooded) prolactin gene: exon iii and flanks
9996	22923	36388	1.67	6.8E-01	AB037768.1	NT	Homo sapiens mRNA for KIAA1345 protein, partial cds
11424	24368	37903	1.77	6.8E-01	AJ276675.1	NT	Stagonospora avenae bgf1 gene for beta-glucosidase, exons 1-4
11424	24368	37904	1.77	6.8E-01	AJ276675.1	NT	Stagonospora avenae bgf1 gene for beta-glucosidase, exons 1-4
11450	24363	37938	1.82	6.8E-01	AF038939.1	NT	Mus musculus zinc finger protein (Pag3) mRNA, complete cds
11450	24363	37939	1.82	6.8E-01	AF038939.1	NT	Mus musculus zinc finger protein (Pag3) mRNA, complete cds
							Mus musculus major histocompatibility complex region NG27, NG28, RPS28, NADH oxidoreductase, NG28, KIFC1, Fas-binding protein, BING1, tapasin, RatGDS-like, KE2, BING4, beta 1,3-galactosyl transferase, and RPS18 genes, complete cds; Sacm21 gene, partial>
11931	24812	38407	1.49	6.8E-01	AF110520.1	NT	Mus musculus major histocompatibility complex region NG27, NG28, RPS28, NADH oxidoreductase, NG28, KIFC1, Fas-binding protein, BING1, tapasin, RatGDS-like, KE2, BING4, beta 1,3-galactosyl transferase, and RPS18 genes, complete cds; Sacm21 gene, partial>
11931	24812	38408	1.49	6.8E-01	AF110520.1	NT	Homo sapiens nuclear factor of kappa light polypeptide gene enhancer in B-cells 1 (NFKB1) gene, complete cds
298	13392	26320	25.45	6.7E-01	AF213884.1	NT	Homo sapiens nuclear factor of kappa light polypeptide gene enhancer in B-cells 1 (NFKB1) gene, complete cds
339	13428	26350	28.03	6.7E-01	AF213884.1	NT	Homo sapiens nuclear factor of kappa light polypeptide gene enhancer in B-cells 1 (NFKB1) gene, complete cds
1927	14951		1.07	6.7E-01	M12132.1	NT	Quail fast skeletal muscle troponin I gene, complete cds

Table 4

Single Exon Probes Expressed in Bone Marrow

Probe SEQ ID NO:	Exon SEQ ID NO:	ORF SEQ ID NO:	Expression Signal	Most Similar (Top) Hit BLAST E Value	Top Hit Accession No.	Top Hit Database Source	Top Hit Descriptor
2167	15173	28192	1.79	6.7E-01	AA451884.1	EST_HUMAN	z12g12.s1 Soares, total_fetus Nb2Hf8_9w Homo sapiens cDNA clone IMAGE:786310 3' similar to contains element TAR1 repetitive element;
2178	15888	28213	2.65	6.7E-01	AF180073.1	NT	Drosophila melanogaster Mst85C gene, complete cds; NMDMC isoform (Nmdmc) gene, complete cds, alternatively spliced; and transcription factor (Relish) gene, complete cds, alternatively spliced
3009	16067	28987	3.96	6.7E-01	8678580	NT	Mus musculus Wiskott-Aldrich syndrome protein (Wasp), mRNA
4481	17506	30395	0.78	6.7E-01	X74421.1	NT	S. tuberosum mRNA for glucose-6-phosphate dehydrogenase
5002	18016	30903	0.97	6.7E-01	AW079110.1	EST_HUMAN	xe9g12.x1 NCL CGAP_Co17 Homo sapiens cDNA clone IMAGE:2574598 3'
5587	18683	31651	0.7	6.7E-01	J04838.1	NT	M.barkeri ATPase alpha and beta subunit (atpA and atpB) genes, complete cds
5587	18683	31652	0.7	6.7E-01	J04838.1	NT	M.barkeri ATPase alpha and beta subunit (atpA and atpB) genes, complete cds
6073	19154	32366	0.93	6.7E-01	AE001486.1	NT	Helicobacter pylori, strain J99 section 47 of 132 of the complete genome
6456	19521	32771	1.26	6.7E-01	8635035	NT	Gallid herpesvirus 2, complete genome
6456	19521	32772	1.26	6.7E-01	8635035	NT	Gallid herpesvirus 2, complete genome
7637	20500	33894	4.57	6.7E-01	AE004608.1	NT	Pseudomonas aeruginosa PA01, section 167 of 529 of the complete genome
7563	20528	33894	0.98	6.7E-01	AE001486.1	NT	Helicobacter pylori, strain J99 section 47 of 132 of the complete genome
10503	23425	37776	0.82	6.7E-01	M34046.1	NT	Human placental protein 14 (PP14) gene, complete cds
11300	24250	37776	2.82	6.7E-01	BF354849.1	EST_HUMAN	CM3-HT0769-010800-197-c3 HT0769 Homo sapiens cDNA
11767	23942	37484	3.23	6.7E-01	O14357	SWISSPROT	N-ACETYL-GLUCOSAMINYL-PHOSPHATIDYLINOSITOL BIOSYNTHETIC PROTEIN GP11
11968	24847	38444	1.82	6.7E-01	AA342521.1	EST_HUMAN	EST48066 Fetal spleen Homo sapiens cDNA 3' end
2509	15512	28536	1.92	6.6E-01	AF075240.1	NT	Homo sapiens SLIT1 protein (SLIT2) mRNA, partial cds
2711	15705	28721	1.4	6.6E-01	AF169339.1	NT	Homo sapiens lens epithelium-derived growth factor gene, alternatively spliced, complete cds
3501	16548	29474	1.41	6.6E-01	4808880	NT	Homo sapiens sema domain, seven thrombospondin repeats (type 1 and type 1-like), transmembrane domain (TM) and short cytoplasmic domain, (semaphorin) 5A (SEMA5A) mRNA
3670	16713	29628	4.37	6.6E-01	V07666.1	NT	C.albicans random DNA marker, 282bp
4136	17168		0.86	6.6E-01	U91328.1	NT	Human hereditary haemochromatosis region, histone 2A-like protein gene, hereditary haemochromatosis (HLA-H) gene, RoRet gene, and sodium phosphate transporter (NPT3) gene, complete cds
5258	18266	31134	1.13	6.6E-01	AL161572.2	NT	Arabidopsis thaliana DNA chromosome 4, contig fragment No. 68
5282	18288	31150	0.95	6.6E-01	Z82002.1	NT	S.pneumoniae popB and popC genes
5282	18288	31151	0.95	6.6E-01	Z82002.1	NT	S.pneumoniae popB and popC genes
6466	19531	32779	3.83	6.6E-01	6690577	NT	Mus musculus kinesin light chain 2 (Klc2), mRNA
7949	20890	34281	3.57	6.6E-01	AV660306.1	EST_HUMAN	AV660508 GLC Homo sapiens cDNA clone GLCGID04 3'
8912	21878	35304	0.58	6.6E-01	AV704700.1	EST_HUMAN	AV704700 ADB Homo sapiens cDNA clone ADBCAF11 5'
10023	22850		1.56	6.6E-01	AL163278.2	NT	Homo sapiens chromosome 21 segment HS21C078

Page 44 of 546
Table 4

Single Exon Probes Expressed in Bone Marrow

Probe SEQ ID NO:	Exon SEQ ID NO:	ORF SEQ ID NO:	Expression Signal	Most Similar (Top) Hit BLAST E Value	Top Hit Accession No.	Top Hit Database Source	Top Hit Descriptor
12746	25395	31768	1.76	6.5E-01	AE004382.1	NT	Vibrio cholerae chromosome II, section 39 of 63 of the complete chromosome
624	13689	26606	0.96	6.5E-01	M75140.1	NT	H. vulgaris Na,K-ATPase alpha subunit mRNA, complete cds
624	13689	26607	0.96	6.5E-01	M75140.1	NT	H. vulgaris Na,K-ATPase alpha subunit mRNA, complete cds
3446	16493	29412	4.63	6.5E-01	AB041225.1	NT	Mus musculus gene for Tob2, complete cds
4311	17340	30220	4.28	6.5E-01	AJ272265.1	NT	Homo sapiens SPP2 gene for secreted phosphoprotein 24 precursor, exons 1-6
5102	16112	30684	3.71	6.5E-01	U28921.1	NT	Phaeodius vulgaris ATPase gamma subunit mRNA, nuclear gene encoding mitochondrial protein, partial cds
5204	18213	31088	1.13	6.5E-01	Z70628.1	NT	H. sapiens mRNA for immunoglobulin heavy chain variable region (8D4-A6, VH4, 4-59DP-71)
6616	25940	31551	2.13	6.5E-01	P18480	SWISSPROT	TRANSCRIPTION REGULATORY PROTEIN SNF5 (SWISNF COMPLEX COMPONENT SNF5)
5802	18894	32077	0.58	6.5E-01	U163249.2	NT	Homo sapiens chromosome 21 segment HS21C049
6887	19939	33234	1.24	6.5E-01	D38348.1	NT	Chicken mRNA for 118-kDa melanosomal matrix protein, complete cds
7841	20788	34163	0.84	6.5E-01	X04750.1	NT	Murine Ig-related lambda(50) gene (exon 1) transcribed selectively in pre-B lymphocytes
7830	20873	34262	0.89	6.5E-01	A1798882.1	EST_HUMAN	wc46a02.x1 NCI_CGAP_P728 Homo sapiens cDNA clone IMAGE:2321642 3'
10197	23122		1.03	6.5E-01	T78904.1	EST_HUMAN	y021b04.e1 Soares fetal liver spleen 1NFLS Homo sapiens cDNA clone IMAGE:108847 3'
10699	23821	37117	2.2	6.5E-01	AF119678.1	NT	Mus musculus small GTP-binding protein RAB26 (Rab26) gene, complete cds
10897	23963	37487	2.19	6.5E-01	H87583.1	EST_HUMAN	yw17706.r1 Soares_placenta_8tc9weeks_2NblHPac9W Homo sapiens cDNA clone IMAGE:252615 5'
11045	24009	37535	2.88	6.5E-01	AA601287.1	EST_HUMAN	no15c07.s1 NCI_CGAP_Phe1 Homo sapiens cDNA clone IMAGE:1100748 3'
11143	24103		3.43	6.5E-01	AU138078.1	EST_HUMAN	AU138078 PLACE1 Homo sapiens cDNA clone PLACE1007610 5'
11825	24806	38309	2.3	6.5E-01	AF014115.1	NT	Plasmodium berghei cytochrome c oxidase subunit III, cytochrome c oxidase subunit I, and cytochrome b genes, mitochondrial genes encoding mitochondrial proteins, complete cds
11870	24848	38445	1.47	6.5E-01	AF146887.1	NT	Fugu rubripes U2 small nuclear ribonucleoprotein auxiliary factor subunit-related protein (U2AF1-RS2), 19 kDa Gcgl adaptor protein adaptin (AP19), and phosphorylase kinase alpha 2 subunit (PHKA2) genes, complete cds; ketch protein (KELCH1) and ketch p2
12130	24969	38604	1.61	6.5E-01	AL161580.2	NT	Arabidopsis thaliana DNA chromosome 4, contig fragment No. 76
12554	25276		2.65	6.5E-01	BE465050.1	EST_HUMAN	hw74a10.x1 NCI_CGAP_Lu24 Homo sapiens cDNA clone IMAGE:3179150 3'
252	13349	26275	10.51	6.4E-01	U48848.1	NT	Drosophila melanogaster 8kd dynein light chain mRNA, complete cds
3470	16516	29437	3.26	6.4E-01	U48854.2	NT	Mus musculus dystroglycan 1 (DAG1) gene, exons 1 and 2 and complete cds
3875	16914	29823	1.34	6.4E-01	AB046827.1	NT	Homo sapiens mRNA for KIAA1607 protein, partial cds
4518	17544	30430	0.89	6.4E-01	Y12488.1	NT	M. musculus wln gene
4519	17544	30431	0.89	6.4E-01	Y12488.1	NT	M. musculus wln gene
5192	18201	31073	1	6.4E-01	H85337.1	EST_HUMAN	y890c08.r1 Soares retina N2b5HR Homo sapiens cDNA clone IMAGE:222066 5'
8960	21928	35353	1.57	6.4E-01	AE001247.1	NT	Trepnema pallidum section 63 of 87 of the complete genome

Page 45 of 546
Table 4

Single Exon Probes Expressed in Bone Marrow

Probe SEQ ID NO:	Exon SEQ ID NO:	ORF SEQ ID NO:	Expression Signal	Most Similar (Top) Hit BLAST E Value	Top Hit Accession No.	Top Hit Database Source	Top Hit Descriptor
10449	23371	36863	6.94	6.4E-01	U82828.1	NT	Homo sapiens ataxia telangiectasia (ATM) gene, complete cds
10484	23388	36879	1.18	6.4E-01	BF670405.1	EST_HUMAN	602150289F1 NIH_MGC_81 Homo sapiens cDNA clone IMAGE:4291126 5'
12668	25342		5.76	6.4E-01	AV758212.1	EST_HUMAN	AV758212 MDS Homo sapiens cDNA clone MDSGCG09 5'
434	13508	26442	3.27	6.3E-01	P05228	SWISSPROT	HISTIDINE-RICH PROTEIN PRECURSOR (CLONE PFHRP-III)
536	13607	28525	2.19	6.3E-01	U32688.1	NT	Haemophilus influenzae Rd section 4 of 163 of the complete genome
2171	15187	28208	3.4	6.3E-01	U81138.1	NT	Shigella flexneri multi-antigen resistance locus
2583	15584	28603	2.72	6.3E-01	U75331.1	NT	Gallus gallus bone morphogenetic protein 1 (BMP-1) mRNA, partial cds
2583	15584	28604	2.72	6.3E-01	U75331.1	NT	Gallus gallus bone morphogenetic protein 1 (BMP-1) mRNA, partial cds
3029	16087		0.8	6.3E-01	Y17275.1	NT	Lycopodium obscurum p69a gene, complete CDS
6182	19257	32480	0.87	6.3E-01	BE093908.1	EST_HUMAN	PMO-BT0757-010500-002-a05 BT0757 Homo sapiens cDNA
6753	19807	33088	1.07	6.3E-01	L27798.1	NT	Streptococcus dysgalactiae (msg) gene, complete cds
6753	19807	33089	1.07	6.3E-01	L27798.1	NT	Streptococcus dysgalactiae (msg) gene, complete cds
8866	21833		3.17	6.3E-01	BE902044.1	EST_HUMAN	601676889F1 NIH_MGC_21 Homo sapiens cDNA clone IMAGE:3958351 5'
9238	22204	35836	0.8	6.3E-01	S62927.1	NT	glycoprotein IIIa (Alu 1 and 3 fusion junction) [human, Genomic Mutant, 300 nt]
8576	22537	35989	0.74	6.3E-01	BF216884.1	EST_HUMAN	601884050F1 NIH_MGC_57 Homo sapiens cDNA clone IMAGE:4102588 5'
9775	22716	36171	2.86	6.3E-01	9827521	NT	Varicella virus, complete genome
9776	22716	36172	2.86	6.3E-01	9827521	NT	Varicella virus, complete genome
10288	23223		0.63	6.3E-01	AE002329.2	NT	Chlamydia muridarum, section 59 of 85 of the complete genome
10792	23713	37215	1.55	6.3E-01	Z73003.1	NT	S. cerevisiae chromosome VII reading frame ORF YGR218w
10896	23815	37322	0.96	6.3E-01	AE000313.1	NT	Escherichia coli K-12 MG1655 section 203 of 400 of the complete genome
11400	24345	37878	1.52	6.3E-01	AA877715.1	EST_HUMAN	nc02408.st NCI_CGAP_Co10 Homo sapiens cDNA clone IMAGE:1161371 3' similar to TR:002916 O02916 HLAR.
11663	24589	38173	6.95	6.3E-01	A804160.1	EST_HUMAN	GM-BT043-080289-048 BT043 Homo sapiens cDNA
11754	24682	38261	1.79	6.3E-01	P47003	SWISSPROT	HYPOTHEITICAL 13.7 KD PROTEIN IN INO1-IDS2 INTERGENIC REGION
11915	24796	38387	2.09	6.3E-01	P36073	SWISSPROT	HYPOTHEITICAL 15.3 KD PROTEIN IN VMA12-APN1 INTERGENIC REGION
12258	25026	31308	5.44	6.3E-01	8810283	NT	Mus musculus keratin complex 2, gene 5g (K12-5g), mRNA
12358	25148		1.81	6.3E-01	AF105277.1	NT	Homo sapiens 3'-phosphoadenosine 5'-phosphosulfate synthetase (PAPSS) mRNA, complete cds
12570	25945		3.19	6.3E-01	X83528.1	NT	C. limicola pscd gene
5970	18055	32255	2.37	6.2E-01	Q10135	SWISSPROT	HYPOTHEITICAL 142.6 KD PROTEIN C23E2.02 IN CHROMOSOME I
7737	20882		2.75	6.2E-01	AF02253.1	NT	Mus musculus calcium-sensing receptor related protein 4 (CaSR-rs4) mRNA, partial cds
7791	25690	34117	1.12	6.2E-01	AL021127.2	NT	Mus musculus chromosome X contigA; putative Magee9 gene, Caltractin, NAD(P) ⁺ steroid dehydrogenase
8644	21612	35034	5.41	6.2E-01	H7255.1	EST_HUMAN	Mus musculus chromosome X contigA; putative Magee9 gene, Caltractin, NAD(P) ⁺ steroid dehydrogenase and Zinc finger protein 185
							ys01608.s1 Soares fetal liver spleen 1NFLS Homo sapiens cDNA clone IMAGE:213542 3'

Table 4

Single Exon Probes Expressed in Bone Marrow

Probe SEQ ID NO:	Exon SEQ ID NO:	ORF SEQ ID NO:	Expression Signal	Most Similar (Top) Hit BLAST E Value	Top Hit Accession No.	Top Hit Database Source	Top Hit Descriptor
9208	22174	36606	0.64	6.2E-01	AF034411.1	NT	Lycopodium esculentum cytosolic Cu,Zn superoxide dismutase (Sod) gene, partial cds, and dehydroquinase dehydroase/ehidmase:NADP oxidoreductase gene, complete cds
9804	21127	34531	1.87	6.2E-01	BE662887.1	EST_HUMAN	601336146F1 NIH_MGC_44 Homo sapiens cDNA clone IMAGE:3690010 5'
9868	22804		2.17	6.2E-01	M24461.1	NT	Human pulmonary surfactant-associated protein SP-B (SFTPB) mRNA, complete cds
10438	23360	36849	7.04	6.2E-01	AL161611.2	NT	Arabidopsis thaliana DNA chromosome 4, contig fragment No. 23
10802	23822	37332	5.32	6.2E-01	P27410	SWISSPROT	NON-STRUCTURAL POLYPEPTIDE [CONTAINS: RNA-DIRECTED RNA POLYMERASE ; THIOL PROTEASE P3C ; HELICASE (2C LIKE PROTEIN); COAT PROTEIN]
10802	23822	37332	5.32	6.2E-01	P27410	SWISSPROT	NON-STRUCTURAL POLYPEPTIDE [CONTAINS: RNA-DIRECTED RNA POLYMERASE ; THIOL PROTEASE P3C ; HELICASE (2C LIKE PROTEIN); COAT PROTEIN]
2404	15411		5.9	6.1E-01	8878076	NT	Mus musculus secreted acidic cysteine rich glycoprotein (Sparc), mRNA
5614	18710	31888	1.3	6.1E-01	M59940.1	NT	Caenorhabditis elegans N2 CehMycD (hlt-1) alternatively spliced genes, complete cds
7053	20075	33382	3.64	6.1E-01	M84733.1	NT	Rat TRPM-2 gene, complete cds
7053	20075	33383	3.54	6.1E-01	M84733.1	NT	Rat TRPM-2 gene, complete cds
7216	20238	33672	0.7	6.1E-01	AW105853.1	EST_HUMAN	cd50H03.x1 NC1_COAP_Ov23 Homo sapiens cDNA clone IMAGE:2597237 3' similar to gb:X12871_ma1 HETEROGENEOUS NUCLEAR RIBONUCLEOPROTEIN A1 (HUMAN);
7312	20283	33824	0.64	6.1E-01	Q63769	SWISSPROT	SUSHI REPEAT-CONTAINING PROTEIN SRPX PRECURSOR (DRS PROTEIN) (DOWN-REGULATED BY V-SRC)
8575	21543	34963	3.68	6.1E-01	AF033535.1	NT	Arabidopsis thaliana putative zinc transporter (ZIP1) mRNA, complete cds
9147	22113	35537	1.17	6.1E-01	11491085	NT	Homo sapiens mitogen-activated protein kinase kinase kinase 4 (MAP4K4), mRNA
9147	22113	35538	1.17	6.1E-01	11491085	NT	Homo sapiens mitogen-activated protein kinase kinase kinase 4 (MAP4K4), mRNA
9770	22711	36165	23.08	6.1E-01	AF236117.1	NT	Homo sapiens G-protein coupled receptor EDG-7 mRNA, complete cds
9770	22711	36166	23.08	6.1E-01	AF236117.1	NT	Homo sapiens G-protein coupled receptor EDG-7 mRNA, complete cds
10202	23127	36613	0.99	6.1E-01	AE004452.1	NT	Pseudomonas aeruginosa PAD1, section 13 of 629 of the complete genome
10406	23328	36812	1.53	6.1E-01	AF119117.1	NT	Homo sapiens dopamine transporter (SLC8A3) gene, complete cds
11308	24256		1.91	6.1E-01	X74507.1	NT	P. salivum mdh mRNA for chloroplast malate dehydrogenase (NADP+)
12041	24916	38510	1.63	6.1E-01	S83182.1	NT	hyaluronan-binding protein-hepatocyte growth factor activator homolog [human, plasma, mRNA, 2408 nt]
12041	24916	38511	1.63	6.1E-01	S83182.1	NT	hyaluronan-binding protein-hepatocyte growth factor activator homolog [human, plasma, mRNA, 2408 nt]
12355	25731	31815	2.77	6.1E-01	AB041350.1	NT	Mus musculus Ccd4a5 mRNA for type IV collagen alpha 5 chain, complete cds
485	13567	26490	1.41	6.0E-01	D87675.1	NT	Homo sapiens DNA for amyloid precursor protein, complete cds
563	13633		2.75	6.0E-01	5802989	NT	Homo sapiens adaptor-related protein complex 3, mu 2 subunit (GLA20), mRNA
1364	14398	27369	1.92	6.0E-01	AF065253.1	NT	Human respiratory syncytial virus strain CI-83-63b attachment protein (G) gene, complete cds
3828	16888	29770	0.9	6.0E-01	AJ233396.1	NT	Viral hemorrhagic septicaemia virus N, P, M, G, N, L genes, French strain 07-71

Table 4

Single Exon Probes Expressed in Bone Marrow

Probe SEQ ID NO:	Exon SEQ ID NO:	ORF SEQ ID NO:	Expression Signal	Most Similar (Top) Hit BLAST E Value	Top Hit Accession No.	Top Hit Database Source	Top Hit Descriptor
4217	17248		1.61	6.0E-01	AF068866.1	NT	Homo sapiens Notch3 (NOTCH3) gene, exons 28, 27, and 28
4279	17308	30187	0.89	6.0E-01	AB025319.1	NT	Yaba monkey tumor virus DNA, BamHI-1 restriction fragment E, M and partial C, partial and complete cds
5353	18468	31327	2.14	6.0E-01	P20288	SWISSPROT	D(2) DOPAMINE RECEPTOR
5514	18814	31547	2.22	6.0E-01	AW139713.1	EST_HUMAN	U1H-B1-est-e-10-0-U1.s1 NCI_CGAP_Sub3 Homo sapiens cDNA clone IMAGE:2718619 3'
6989	19746	33022	2.68	6.0E-01	U38813.1	NT	Musca domestica insecticide-susceptible strain voltage-sensitive sodium channel mRNA, complete cds
6818	19872	33181	0.87	6.0E-01	Q04912	SWISSPROT	MACROPHAGE-STIMULATING PROTEIN RECEPTOR PRECURSOR (MSP RECEPTOR) (P185-RON)
6988	20211	33539	0.78	6.0E-01	L10234.1	NT	(CDW136) (CD136 ANTIGEN)
6988	20211	33540	0.78	6.0E-01	L10234.1	NT	Strongylocentrotus purpuratus kinesin light chain isoform 2 mRNA, complete cds
7577	20539	33898	5.51	6.0E-01	AJ277681.1	NT	Strongylocentrotus purpuratus kinesin light chain isoform 2 mRNA, complete cds
8461	21430	34847	4.55	6.0E-01	P02835	SWISSPROT	Homo sapiens partial LMO1 gene for LIM domain only 1 protein, exon 1
8461	21430	34848	4.55	6.0E-01	P02835	SWISSPROT	SEGMENTATION PROTEIN FUSHI TARAZU
10182	23107	36588	1.84	6.0E-01	AB008193.1	NT	Homo sapiens genes for leukotriene B4 receptor BLT2, leukotriene B4 receptor BLT1, complete cds
10336	23558		1.66	6.0E-01	Q01487	SWISSPROT	PEROXISOMAL MEMBRANE PROTEIN PERO (PEROXIN-3)
10749	23871		0.48	6.0E-01	BE837778.1	EST_HUMAN	RC2-FN0094-180700-017-408 FN0094 Homo sapiens cDNA
11878	24760	38345	2.79	6.0E-01	AI420623.1	EST_HUMAN	if0807.X1 NCI_CGAP_Pr28 Homo sapiens cDNA clone IMAGE:2095621 3'
12638	25322	31798	1.87	6.0E-01	11421683	NT	Homo sapiens nuclear factor (erythroid-derived 2)-like 3 (NFE2L3), mRNA
12731	25384		2.78	6.0E-01	AA700087.1	EST_HUMAN	z98g05.s1 Soares_fetal_liver_spleen_INFLS_S1 Homo sapiens cDNA clone IMAGE:482776 3'
12918	25777	31522	4.71	6.0E-01	9056303	NT	Mus musculus cGMP-inhibited phosphodiesterase (Pde3a), mRNA
12947	25715		3.4	6.0E-01	BE157617.1	EST_HUMAN	RC1-HT0375-030500-015-c03 HT0375 Homo sapiens cDNA
1002	14053	27009	0.97	5.9E-01	U32701.1	NT	Haemophilus influenzae Rd section 16 of 163 of the complete genome
3283	16337	29256	4.95	5.9E-01	AL163287.2	NT	Homo sapiens chromosome 21 segment HS21C087
3283	16337	29257	4.95	5.9E-01	AL163287.2	NT	Homo sapiens chromosome 21 segment HS21C087
4250	17279		4.09	5.9E-01	AF162758.1	NT	Rattus norvegicus ceneixin 2 mRNA, partial cds
6909	19887	32943	1.45	5.9E-01	AF068440.2	NT	Homo sapiens low density lipoprotein receptor-related protein II (LRP2) gene, exon 1 and partial cds
7481	20447	33803	2.44	5.9E-01	AB023486.1	NT	Homo sapiens gene for histamine H2 receptor, promoter region and complete cds
6332	21301	34718	0.48	5.9E-01	D90911.1	NT	Synechocystis sp. PCC6803 complete genome, 13/27, 1578593-1719843
8985	21951	35375	0.48	5.9E-01	D12922.1	NT	Legionella pneumophila gene for iron superoxide dismutase, complete cds
9901	22853	36314	0.89	5.9E-01	AF063204.2	NT	Chlamydia trachomatis strain K/UW/31/Cx major outer membrane protein (omp1) gene, complete cds

Table 4

Single Exon Probes Expressed in Bone Marrow

Probe SEQ ID NO:	Exon SEQ ID NO:	ORF SEQ ID NO:	Expression Signal	Most Similar (Top) Hit BLAST E Value	Top Hit Accession No.	Top Hit Database Source	Top Hit Descriptor
10273	23188		0.66	5.8E-01	P08463	SWISSPROT	E6 PROTEIN
10548	23470	36865	1.19	5.8E-01	P56284	SWISSPROT	VASCULAR ENDOTHELIAL-CADHERIN PRECURSOR (VE-CADHERIN) (CADHERIN-5)
11031	23895	37523	2.46	5.8E-01	Q9X013	SWISSPROT	THYMIDYLATE KINASE (DTMP KINASE)
11037	24001	37526	49.8	5.8E-01	AF187944.1	NT	Xenopus laevis receptor protein tyrosine phosphatase delta (XPTP-D) mRNA, complete cds
11309	24259	37785	2.49	5.8E-01	AW937176.1	EST_HUMAN	PM1-DT0041-180100-002-h03 DT0041 Homo sapiens cDNA
11526	24467	38021	1.53	5.8E-01	AF064628.1	NT	Mus spretus strain SPRET/EJ CD48 antigen (Cd48) gene, partial cds
12297	25108	31838	2.43	5.8E-01	L42320.1	NT	Oryctolagus cuniculus alpha 1 anti-trypsin (alpha 1 AT) gene, promoter region
12641	25264		2.86	5.8E-01	AB017705.1	NT	Aspergillus oryzae pyrG gene for orotidine-5'-phosphate decarboxylase, complete cds
12740	25390		6.24	5.8E-01	P34828	SWISSPROT	MICROTUBULE-ASSOCIATED PROTEIN 1A [CONTAINS: MAP1 LIGHT CHAIN LC2]
1824	14948	27944	1.28	5.8E-01	P40472	SWISSPROT	SIM1 PROTEIN
2571	15572	28592	1	5.8E-01	7305230	NT	Mus musculus low density lipoprotein B (Ldlb), mRNA
4009	17048	28654	1	5.8E-01	BF685738.1	EST_HUMAN	601852474F1 NIH_MGC_56 Homo sapiens cDNA clone IMAGE:4076131 5'
4542	17965	30482	4.23	5.8E-01	AB009077.1	NT	Vigna radiata mRNA for protein pyrophosphatase, complete cds
5448	18550		0.84	5.8E-01	AE002152.1	NT	Ureaplasma urealyticum section 53 of 59 of the complete genome
5609	18705	31862	0.75	5.8E-01	Q10699	SWISSPROT	POTENTIAL 5'-3' EXONUCLEASE
6308	18378	32618	1.82	5.8E-01	D76659.1	EST_HUMAN	HUM500E088 Human placenta polyA+ (TFijwara) Homo sapiens cDNA clone GEN-500E08 5'
6445	19510	32760	0.73	5.8E-01	D50601.1	NT	Shigella sonnei DNA for 28 ORFs, complete cds
6885	20208		2.55	5.8E-01	S65091.1	NT	cyclic AMP-regulated phosphoprotein [rats, mRNA, 1030 nt]
8219	21188		2.63	5.8E-01	H41571.1	EST_HUMAN	Y81503.s1 Soares adult brain N256HB50Y Homo sapiens cDNA clone IMAGE:176767 3' similar to gb:S78187 M-PHASE INDUCER PHOSPHATASE 2 (HUMAN);
8423	21362	34802	0.59	5.8E-01	A1280051.1	EST_HUMAN	q85d10.x1 Soares_NFL_T_GBC_S1 Homo sapiens cDNA clone IMAGE:1853779 3'
8423	21362	34803	0.59	5.8E-01	A1280051.1	EST_HUMAN	q85d10.x1 Soares_NFL_T_GBC_S1 Homo sapiens cDNA clone IMAGE:1853779 3'
8532	21500	34916	2.57	5.8E-01	P14328	SWISSPROT	SPORE COAT PROTEIN SP86
8532	21500	34917	2.57	5.8E-01	P14328	SWISSPROT	SPORE COAT PROTEIN SP86
9243	22209	35940	9.77	5.8E-01	A1270774.1	NT	Homo sapiens partial TCF-4 gene for T-cell transcription factor-4, exons 6-11
9243	22288	35718	1.02	5.8E-01	Q27368	SWISSPROT	TRANSCRIPTION FACTOR E2F
9324	22289	35719	0.46	5.8E-01	Q20471	SWISSPROT	POTATIVE CASEIN KINASE I F46F2.2 IN CHROMOSOME X
9854	22981		0.81	5.8E-01	BF031606.1	EST_HUMAN	601557774F1 NIH_MGC_58 Homo sapiens cDNA clone IMAGE:3827288 5'
11334	24284	37808	0.9	5.8E-01	A1243213.1	NT	Homo sapiens partial 5-HT4 receptor gene, exons 2 to 5
11373	24320		2.69	5.8E-01	BF700062.1	EST_HUMAN	602127577F1 NIH_MGC_56 Homo sapiens cDNA clone IMAGE:4284403 5'
11476	24419		1.61	5.8E-01	BF700062.1	EST_HUMAN	602127577F1 NIH_MGC_56 Homo sapiens cDNA clone IMAGE:4284403 5'
1492	14625	27498	1.11	5.7E-01	P06727	SWISSPROT	APOLIPOPROTEIN A-IV PRECURSOR (APO-AIV)
1492	14625	27497	1.11	5.7E-01	P06727	SWISSPROT	APOLIPOPROTEIN A-IV PRECURSOR (APO-AIV)
3056	16113		0.77	5.7E-01	6755253	NT	Mus musculus plasmacytoma variant translocation 1 (Pvt1), mRNA

Page 49 of 546
Table 4

Single Exon Probes Expressed in Bone Marrow

Probe SEQ ID NO:	Exon SEQ ID NO:	ORF SEQ ID NO:	Expression Signal	Most Similar (Top) Hit BLAST E Value	Top Hit Accession No.	Top Hit Database Source	Top Hit Descriptor
3237	16292	29214	1.38	5.7E-01	Q9WTJ2	SWISSPROT	PUTATIVE TRANSCRIPTION FACTOR OVO-LIKE 1 (MOV01) (MOV01A)
3516	16561		2.79	5.7E-01	AB033503.1	NT	Populus euphratica pease-2 mRNA for 1-aminocyclopropane-1-carboxylate synthase, complete cds
3922	16962	29876	1.05	5.7E-01	AF011581.1	NT	Homo sapiens T cell receptor beta chain (BV6S72-BJ1S1) mRNA, partial cds
5213	18222	31087	11.22	5.7E-01	4505050	NT	Homo sapiens lymphocyte antigen 6 complex, locus H (LY6H) mRNA
6490	19555	32805	4.36	5.7E-01	BF035413.1	EST_HUMAN	601454962F1 NIH_MGC 66 Homo sapiens cDNA clone IMAGE:3838590 5'
6869	19922	33218	0.82	5.7E-01	AA194201.1	EST_HUMAN	z38c08.1 Soares_NHMPu_S1 Homo sapiens cDNA clone IMAGE:865674 5'
7042	18374	31282	1.3	5.7E-01	AL111440.1	NT	Bordetella pertussis strain T4 cDNA library under conditions of nitrogen deprivation
8041	20978	34374	2.13	5.7E-01	P00373	SWISSPROT	PYRROLINE-5-CARBOXYLATE REDUCTASE (P5CR) (P5C REDUCTASE)
8303	21272		0.51	5.7E-01	AJ251835.1	NT	Mus musculus Kcpq1, Ltrpc5, Mash2, Tapa-1, Tsc4 and Tsc6 genes, alternative transcripts
8723	21691		0.82	5.7E-01	AI065061.1	EST_HUMAN	HA0885 Human fetal liver cDNA library Homo sapiens cDNA
10159	23084	36580	1.22	5.7E-01	AL161532.2	NT	Arabidopsis thaliana DNA chromosome 4, contig fragment No. 32
10159	23084	36591	1.22	5.7E-01	AL161532.2	NT	Arabidopsis thaliana DNA chromosome 4, contig fragment No. 32
10940	23860	37376	0.75	5.7E-01	BF549882.1	EST_HUMAN	602067712F1 NIH_MGC 58 Homo sapiens cDNA clone IMAGE:4066810 5'
12262	25078		1.49	5.7E-01	BE715051.1	EST_HUMAN	MR3-HT0736-180700-003-e02 HT0736 Homo sapiens cDNA
1889	14914	27907	1.6	5.6E-01	AF097732.1	NT	Homo sapiens caspase recruitment domain-containing protein (BCL10) gene, complete cds
1889	14914	27908	1.6	5.6E-01	AF097732.1	NT	Homo sapiens caspase recruitment domain-containing protein (BCL10) gene, complete cds
3376	16426	29351	1.53	5.6E-01	AB018283.2	NT	Homo sapiens mRNA for KIAA0740 protein, partial cds
3376	16426	29352	1.53	5.6E-01	AB018283.2	NT	Homo sapiens mRNA for KIAA0740 protein, partial cds
4268	17297	30174	0.77	5.6E-01	D83135.1	NT	Chicken TBP gene, exon8, complete cds
5205	18214	31089	0.93	5.6E-01	BF032377.1	EST_HUMAN	601452855F1 NIH_MGC 66 Homo sapiens cDNA clone IMAGE:3856717 5'
9155	22121	35550	14.86	5.6E-01	AV684703.1	EST_HUMAN	AV684703 GKC Homo sapiens cDNA clone GKCF5F05 5'
9155	22121	35551	14.86	5.6E-01	AV684703.1	EST_HUMAN	AV684703 GKC Homo sapiens cDNA clone GKCF5F05 5'
9730	22758	36211	1.23	5.6E-01	AB038782.1	NT	Homo sapiens MUC3A gene for intestinal mucin, partial cds
12163	25011		3.4	5.6E-01	BE888280.1	EST_HUMAN	601514007F1 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:3915457 5'
12270	25082	38178	1.73	5.6E-01	AA493535.1	EST_HUMAN	ng76g10.a1 NCI_QGAP_Pr6 Homo sapiens cDNA clone IMAGE:3940674 similar to contains element PTR7 repetitive element;
12636	18341	31290	1.51	5.6E-01	AL161501.2	NT	Arabidopsis thaliana DNA chromosome 4, contig fragment No. 13
12662	26340		3.05	5.6E-01	P50505	SWISSPROT	HIGH AFFINITY POTASSIUM TRANSPORTER
13060	25596		4.95	5.6E-01	BF573829.1	EST_HUMAN	602132029F1 NIH_MGC 81 Homo sapiens cDNA clone IMAGE:4271334 5'
13110	25630		1.33	5.6E-01	AA663881.1	EST_HUMAN	ae74804.s1 Stralagene schizo brain S11 Homo sapiens cDNA clone IMAGE:969871 3'
1216	14254	27212	0.82	5.6E-01	8393912	NT	Rattus norvegicus Protonyl Coenzyme A carboxylase, beta polypeptide (Pccb), mRNA
2712	15706	28722	5.31	5.5E-01	P03341	SWISSPROT	GAG POLYPROTEIN [CONTAINS: INNER COAT PROTEIN P12; CORE PROTEIN P16; CORE SHELL PROTEIN P30; NUCLEOPROTEIN P10]

Table 4

Single Exon Probes Expressed In Bone Marrow

Probe SEQ ID NO:	Exon SEQ ID NO:	ORF SEQ ID NO:	Expression Signal	Most Similar (Top) Hit BLAST E Value	Top Hit Accession No.	Top Hit Database Source	Top Hit Descriptor
2712	15706	28723	5.31	5.5E-01	P03341	SWISSPROT	GAG POLYPEPTIDE [CONTAINS: INNER COAT PROTEIN P12; CORE PROTEIN P16; CORE SHELL PROTEIN P30; NUCLEOPROTEIN P10]
2829	15987	28908	0.78	5.5E-01	5902086	NT	Homo sapiens superkiller viral activity 2 (S. cerevisiae homolog)-like (SKN2), mRNA
3079	16136		1.48	5.5E-01	H48219.1	EST_HUMAN	yo18e10.s1 Soares adult brain N2b5HB5Y Homo sapiens cDNA clone IMAGE:178268 3'
3248	16303	29227	4.75	5.5E-01	AF227240.1	NT	Rabbit oral papillomavirus, complete genome
3704	16747	29861	2.24	5.5E-01	P48755	SWISSPROT	FOS-RELATED ANTIGEN-1
5209	18218	31094	1.06	5.5E-01	U69097.1	NT	Bos taurus MHC class II beta-chain BoLA-DIB1 gene, partial cds
7467	20433	33789	0.58	5.5E-01	AF030001.1	NT	Mus musculus major histocompatibility locus class III region: butyrophilin-like protein gene, partial cds; Notch4, PBX2, RAGE, lysophosphatidic acid acyl transferase-alpha, palmitoyl-protein thioesterase 2 (PPT2), CREB-RP, and tenascin X (TNX) genes, complete
7467	20433	33780	0.58	5.5E-01	AF030001.1	NT	Mus musculus major histocompatibility locus class III region: butyrophilin-like protein gene, partial cds; Notch4, PBX2, RAGE, lysophosphatidic acid acyl transferase-alpha, palmitoyl-protein thioesterase 2 (PPT2), CREB-RP, and tenascin X (TNX) genes, complete
7504	20469		0.67	5.5E-01	AB015596.1	NT	Carassius auratus gene for gonadotropin II beta subunit, complete cds
8782	21759	35181	0.66	5.5E-01	A1791786.1	EST_HUMAN	alpha2501.y5 NCI CGAP_Lu5 Homo sapiens cDNA clone IMAGE:1602336 5'
10125	23051		0.69	5.5E-01	U88416.1	NT	Crimean-Congo hemorrhagic fever virus strain SPU 415/85 nucleoprotein gene, complete cds
10743	23665	37160	0.87	5.5E-01	T05047.1	EST_HUMAN	EST02936 Fetal brain, Stratagene (cat#826206) Homo sapiens cDNA clone HFBQ336
146	13249	26178	9.02	5.4E-01	7667266	NT	Homo sapiens KIAA0929 protein Msc2 interacting nuclear target (MINT) homolog (KIAA0929), mRNA
146	13249	26179	9.02	5.4E-01	7667266	NT	Homo sapiens KIAA0929 protein Msc2 interacting nuclear target (MINT) homolog (KIAA0929), mRNA
587	13655	26669	1.34	5.4E-01	AF232008.1	NT	Pseudomonas syringae pv. tomato strain DC3000 AvrE (avrE), HrpW (hrpW), and Gsta (gsta) genes, complete cds; and unknown genes
587	13655	26670	1.34	5.4E-01	AF232008.1	NT	Pseudomonas syringae pv. tomato strain DC3000 AvrE (avrE), HrpW (hrpW), and Gsta (gsta) genes, complete cds; and unknown genes
1278	14911	27272	2.89	5.4E-01	AW896037.1	EST_HUMAN	QV4-NIN040-070400-160-c04 NN0040 Homo sapiens cDNA
2116	15133		2.81	5.4E-01	AE002247.2	NT	Chlamydia pneumoniae AR39, section 74 of 94 of the complete genome
2265	15279	28304	2.26	5.4E-01	AJ276882.1	NT	Drosophila melanogaster mRNA for 15.15' beta carboxylate dioxygenase (beta-diox gene)
5740	18634	32014	0.91	5.4E-01	AW842327.1	EST_HUMAN	PM2-CN0030-030200-003-c10 CN0030 Homo sapiens cDNA
6315	18986	32628	0.8	5.4E-01	AB025017.1	NT	Rattus norvegicus gene for TIS11, complete cds
7228	20248	33582	0.77	5.4E-01	BE868592.2	EST_HUMAN	601680276F1 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:3006080 3'
7558	20521	33877	0.76	5.4E-01	Z21619.1	NT	S.cerevisiae RIB3 gene encoding DBP synthase
7558	20521	33878	0.76	5.4E-01	Z21619.1	NT	S.cerevisiae RIB3 gene encoding DBP synthase

Page 51 of 546
Table 4
Single Exon Probes Expressed in Bone Marrow

Probe SEQ ID NO:	Exon SEQ ID NO:	ORF SEQ ID NO:	Expression Signal	Most Similar (Top) Hit BLAST E Value	Top Hit Accession No.	Top Hit Database Source	Top Hit Descriptor
7580	20523	33881	1.78	5.4E-01	Q64428	SWISSPROT	MITOCHONDRIAL TRIFUNCTIONAL ENZYME ALPHA SUBUNIT PRECURSOR (TP-ALPHA)
10349	23273		1.93	5.4E-01	BF572838.1	EST_HUMAN	[INCLUDES: LONG-CHAIN ENOYL-COA HYDRATASE; LONG CHAIN 3-HYDROXYACYL-COA DEHYDROGENASE]
11414	24358	37893	2.19	5.4E-01	P36888	SWISSPROT	602076545F1 NIH_MGC_62 Homo sapiens cDNA clone IMAGE:4243680 5'
11652	24589	38159	1.82	5.4E-01	AW373694.1	EST_HUMAN	NITRATE REDUCTASE [NADPH] (NIR)
11941	24821	38416	3.29	5.4E-01	Q60675	SWISSPROT	QV4-BT0536-271289-068-H04 BT0536 Homo sapiens cDNA
11941	24821	38417	3.29	5.4E-01	Q60675	SWISSPROT	LAMININ ALPHA-2 CHAIN PRECURSOR (LAMININ M CHAIN) (MEROSIN HEAVY CHAIN)
12215	25054		3.88	5.4E-01	A1858398.1	EST_HUMAN	LAMININ ALPHA-2 CHAIN PRECURSOR (LAMININ M CHAIN) (MEROSIN HEAVY CHAIN)
517	13588	28508	1.86	5.3E-01	AF019413.1	NT	LAMININ ALPHA-2 CHAIN PRECURSOR (LAMININ M CHAIN) (MEROSIN HEAVY CHAIN)
2150	15168	28182	0.97	5.3E-01	AF113919.1	NT	w37604.x1 NCI_CGAP_UH1 Homo sapiens cDNA clone IMAGE:2427128 3' similar to gb:M13452 LAMIN A (HUMAN);
2160	15168	28183	0.97	5.3E-01	AF113919.1	NT	Homo sapiens HLA class III region containing tenascin X (tenascin-X) gene, partial cds; cytochrome P450 21- hydroxylase (CYP21B), complement component C4 (C4B) G11, helicase (SKI2W), RD, complement factor B (Bf), and complement component C2 (C2) genes;>
2794	15786	28803	8.62	5.3E-01	4506328	NT	Brassicaceae var. capitata phospholipase D2 (PLD2) gene, complete cds
2794	15786	28804	8.62	5.3E-01	4506328	NT	Brassicaceae var. capitata phospholipase D2 (PLD2) gene, complete cds
3257	16311	29232	3.25	5.3E-01	AF087658.1	NT	Homo sapiens protein tyrosine phosphatase, receptor-type, zeta polypeptide 1 (PTPRZ1) mRNA
4239	17268		1.33	5.3E-01	U59687.1	NT	Homo sapiens protein tyrosine phosphatase, receptor-type, zeta polypeptide 1 (PTPRZ1) mRNA
5533	18631	31569	2.06	5.3E-01	A1820921.1	EST_HUMAN	Homo sapiens secreted C-type lectin precursor (LIS1CL) gene, complete cds
5533	18631	31570	2.06	5.3E-01	A1820921.1	EST_HUMAN	Mycoplasma genitalium section 9 of 51 of the complete genome
5633	18729	31860	0.76	5.3E-01	AA193672.1	EST_HUMAN	zu42h12.j5 Soares ovary tumor NbHOT Homo sapiens cDNA clone IMAGE:740711 5'
5633	18729	31891	0.76	5.3E-01	AA193672.1	EST_HUMAN	zu42h12.j5 Soares ovary tumor NbHOT Homo sapiens cDNA clone IMAGE:740711 5'
5729	18823	32003	2	5.3E-01	BE645620.1	EST_HUMAN	zu42g09.r1 Soares NIHIMPu_S1 Homo sapiens cDNA clone IMAGE:668112 5'
5729	18823	32004	2	5.3E-01	BE645620.1	EST_HUMAN	zu42g09.r1 Soares NIHIMPu_S1 Homo sapiens cDNA clone IMAGE:668112 5'
9255	22221		1.94	5.3E-01	L01050.2	NT	7673c12.x1 NCI_CGAP_P28 Homo sapiens cDNA clone IMAGE:3288118 3' similar to gb:J02783
9307	22272	35703	0.76	5.3E-01	BF433958.1	EST_HUMAN	PROTEIN DISULFIDE ISOMERASE PRECURSOR (HUMAN);
9307	22272	35704	0.76	5.3E-01	BF433958.1	EST_HUMAN	7673c12.x1 NCI_CGAP_P28 Homo sapiens cDNA clone IMAGE:3288118 3' similar to gb:J02783
							PROTEIN DISULFIDE ISOMERASE PRECURSOR (HUMAN);
							Roridula gorgonites ribulose 1,5-bisphosphate carboxylase (rbcl) gene, partial cds; chloroplast gene for chloroplast product
							7q71c12.x1 NCI_CGAP_Lu24 Homo sapiens cDNA clone IMAGE: 3' similar to contains element MER29 repetitive element;
							7q71c12.x1 NCI_CGAP_Lu24 Homo sapiens cDNA clone IMAGE: 3' similar to contains element MER29 repetitive element;

Table 4

Single Exon Probes Expressed in Bone Marrow

Probe SEQ ID NO:	Exon SEQ ID NO:	ORF SEQ ID NO:	Expression Signal	Most Similar (Top) Hit BLAST E Value	Top Hit Accession No.	Top Hit Database Source	Top Hit Descriptor
10572	23494	36886	0.74	5.3E-01	A1854210.1	EST_HUMAN	w84b02.x1 NCL_CGAP_Med15 Homo sapiens cDNA clone IMAGE:2551275 3' similar to SW:COXA_HUMAN P20874 CYTOCHROME C OXIDASE POLYPEPTIDE VA PRECURSOR ;
10977	23897	37410	0.7	5.3E-01	11428833	NT	Homo sapiens nucleoporin 214kD (CAIN) (NUP214), mRNA
11886	24767	38354	5.19	5.3E-01	BE568291.1	EST_HUMAN	601333857F1 NIH_MGC_53 Homo sapiens cDNA clone IMAGE:3882168 5'
12144	25775		4.97	5.3E-01	AA916053.1	EST_HUMAN	og30605.s1 NCL_CGAP_Br7 Homo sapiens cDNA clone IMAGE:1441376 3' similar to gb:J02611
817	13875	26823	16.24	5.2E-01	L20770.1	NT	APOLIPROTEIN D PRECURSOR (HUMAN); Drosophila melanogaster helix-loop-helix mRNA, complete cds
1168	14208	27163	7.88	5.2E-01	Q9WV30	SWISSPROT	NUCLEAR FACTOR OF ACTIVATED T CELLS 5 (T CELL TRANSCRIPTION FACTOR NFAT5) (NF-A15)
1165	14235	27180	3.01	5.2E-01	AF224492.1	NT	(REL DOMAIN-CONTAINING TRANSCRIPTION FACTOR NFAT5)
1803	14927		3.19	5.2E-01	AL163285.2	NT	Homo sapiens phospholipid scramblase 1 gene, complete cds
2156	15172	28191	2.36	5.2E-01	AB018283.2	NT	Homo sapiens chromosome 21 segment HS21C085
3136	16193	29102	1.57	5.2E-01	U65942.1	NT	Homo sapiens mRNA for KIAA0740 protein, partial cds
3251	16306		1.14	5.2E-01	D73443.1	NT	Chlamydomonas abortus strain S283 POM91A and POM90A precursor, genes, complete cds
3416	16494		1.39	5.2E-01	AL116780.1	NT	Azotobacter vinelandii lsd gene for isocitrate dehydrogenase, complete cds
3458	16504	29424	2.72	5.2E-01	AA984165.1	EST_HUMAN	Botrytis cinerea strain T4 cDNA library under conditions of nitrogen deprivation
3648	16681		1	5.2E-01	AF020289.1	NT	am77g05.s1 Strabagene schizo brain S11 Homo sapiens cDNA clone IMAGE:1616504 3'
5180	18189		1.04	5.2E-01	AL163281.2	NT	Medicago sativa chloroplast malate dehydrogenase precursor (p1mdh) mRNA, nuclear gene encoding chloroplast protein, complete cds
5736	18830	32008	1.02	5.2E-01	AA284261.1	EST_HUMAN	Homo sapiens chromosome 21 segment HS21C081
10088	25700	36489	0.84	5.2E-01	X02218.1	NT	zc44409.T7 Soares_senscent_fibroblasts_NbHSF Homo sapiens cDNA clone IMAGE:325169 3'
10088	25700	36490	0.84	5.2E-01	X02218.1	NT	Chicken duplicated genes for histone H2A, H4 and a histone H3 gene
10292	23217	36701	0.52	5.2E-01	AA194518.1	EST_HUMAN	Chicken duplicated genes for histone H2A, H4 and a histone H3 gene
10387	23308	36787	1.76	5.2E-01	AF143952.2	NT	zq05b08.J1 Strabagene muscle 937209 Homo sapiens cDNA clone IMAGE:628783 5'
13031	25578		6.62	5.2E-01	P18516	SWISSPROT	Homo sapiens PELOTA (PELOTA) gene, complete cds
619	13684	26602	2.34	5.1E-01	M59509.1	NT	RETINOIC ACID RECEPTOR GAMMA (RAR-GAMMA) (RETINOIC ACID RECEPTOR DELTA) (RAR-DELTA)
649	13715	26636	3.53	5.1E-01	AJ233944.1	NT	Human adrenodoxin reductase gene, exons 3 to 12
649	13715	26637	3.53	5.1E-01	AJ233944.1	NT	Polyangium vitalinum (strain Pt vt1) 16S rRNA gene
1660	14692		1.28	5.1E-01	X87885.1	NT	Polyangium vitalinum (strain Pt vt1) 16S rRNA gene
4103	17137	30032	4.81	5.1E-01	AJ859495.1	EST_HUMAN	R.norvegicus mRNA for mammalian fusca protein
4216	17245	30130	2.99	5.1E-01	P86380	SWISSPROT	w39612.x1 NCL_CGAP_UH1 Homo sapiens cDNA clone IMAGE:2427263 3'
6348	19417	32658	0.57	5.1E-01	BE541068.1	EST_HUMAN	TRANSCRIPTION-REPAIR COUPLING FACTOR (TRCF)
6407	19475		0.83	5.1E-01	AV712326.1	EST_HUMAN	601063606F1 NIH_MGC_10 Homo sapiens cDNA clone IMAGE:3450000 5'
							AV712326 DCA Homo sapiens cDNA clone DCAAU07 5'

Table 4

Single Exon Probes Expressed in Bone Marrow

Probe SEQ ID NO:	Exon SEQ ID NO:	ORF SEQ ID NO:	Expression Signal	Most Similar (Top) Hit BLAST E Value	Top Hit Accession No.	Top Hit Database Source	Top Hit Descriptor
7102	20036	33339	1.52	5.1E-01	R80873.1	EST_HUMAN	y94a09.s1 Soares placenta Nb2-IP Homo sapiens cDNA clone IMAGE:146872 3'
8918	21884	35309	0.82	5.1E-01	AW806881.1	EST_HUMAN	QV4-ST0023-160400-172-e01 ST0023 Homo sapiens cDNA
8918	21884	35310	0.82	5.1E-01	AW806881.1	EST_HUMAN	QV4-ST0023-160400-172-e01 ST0023 Homo sapiens cDNA
10043	22970	36437	4.3	5.1E-01	J05412.1	NT	Human regenerating protein (reg) gene, complete cds
10046	22973	36440	3.2	5.1E-01	W22302.1	EST_HUMAN	56B1 Human retina cDNA Tsp508I-cleaved sublibrary Homo sapiens cDNA not directional
10521	23443	36941	0.89	5.1E-01	M94579.1	NT	Human carboxyl ester lipase (CEL) gene, complete cds
12264	25088	38174	2.09	5.1E-01	BF540777.1	EST_HUMAN	602067471F1 NIH_MGC_58 Homo sapiens cDNA clone IMAGE:4066744 5'
12368	25709		2.47	5.1E-01	BF030207.1	EST_HUMAN	601566883F1 NIH_MGC_58 Homo sapiens cDNA clone IMAGE:3826767 5'
12811	25310		3.82	5.1E-01	BF439982.1	EST_HUMAN	ncs51f10.x1 NCI_CGAP_Bm23 Homo sapiens cDNA clone IMAGE:3406218 3' similar to contains element TAR1 repetitive element:
2144	15161	28176	0.97	5.0E-01	4885552	NT	Homo sapiens postmitotic segregation increased 2-like 9 (PMS2L9), mRNA
2144	15161	28177	0.97	5.0E-01	4885552	NT	Homo sapiens postmitotic segregation increased 2-like 9 (PMS2L9), mRNA
2154	15170	28187	2.39	5.0E-01	AF008210.1	NT	Buchnera aphidicola genomic fragment containing (chaperone Hsp60) groEL, DNA biosynthesis initiating protein (dnaA), ATP operon (atpCDGAHFEb), and putative chromosome replication protein (gidA) genes, complete cds; and termination factor Rho (rho) gene>
2154	15170	28188	2.39	5.0E-01	AF008210.1	NT	Buchnera aphidicola genomic fragment containing (chaperone Hsp60) groEL, DNA biosynthesis initiating protein (dnaA), ATP operon (atpCDGAHFEb), and putative chromosome replication protein (gidA) genes, complete cds; and termination factor Rho (rho) gene>
2172	15186		0.91	5.0E-01	AL161533.2	NT	Arabidopsis thaliana DNA chromosome 4, contig fragment No. 33
3761	16802	29714	0.9	5.0E-01	U55574.1	NT	Mus musculus anti-DNA immunoglobulin light chain IgM mRNA, antibody 363p.138, partial cds
3842	16882	29786	1	5.0E-01	L38483.1	NT	Rattus norvegicus jagged protein mRNA, complete cds
3887	16927	29835	2.74	5.0E-01	AB033010.1	NT	Rattus norvegicus mRNA for KIAA1184 protein, partial cds
6901	19855		0.84	5.0E-01	BF576189.1	EST_HUMAN	Homo sapiens mRNA for KIAA1184 protein, partial cds
7926	20869	34256	0.86	5.0E-01	AL161549.2	NT	602132842F1 NIH_MGC_81 Homo sapiens cDNA clone IMAGE:4271839 5'
7926	20869	34257	0.86	5.0E-01	AL161549.2	NT	Arabidopsis thaliana DNA chromosome 4, contig fragment No. 49
8875	21842		1.87	5.0E-01	M92304.1	NT	Arabidopsis thaliana DNA chromosome 4, contig fragment No. 49
9018	21884	35404	0.59	5.0E-01	BF107848.1	EST_HUMAN	Xenopus laevis smooth muscle beta-tropomyosin mRNA, complete cds
9813	21136	34538	3.16	5.0E-01	BF317212.1	EST_HUMAN	601823850R1 NIH_MGC_79 Homo sapiens cDNA clone IMAGE:4043465 3'
							601903871F1 NIH_MGC_19 Homo sapiens cDNA clone IMAGE:4136832 5'
							GLYCAGEN DEBRANCHING ENZYME (GLYCAGEN DEBRANCHER) [INCLUDES: 4-ALPHA-GLUCANOTRANSFERASE (OLIGO-1,4-1,4-GLUCANTRANSFERASE); AMYLO-1,6-GLUCOSIDASE (DEXTRIN 6-ALPHA-D-GLUCOSIDASE)]
9883	22910	36376	1.31	5.0E-01	P35573	SW/ISSPROT	GLYCAGEN DEBRANCHING ENZYME (GLYCAGEN DEBRANCHER) [INCLUDES: 4-ALPHA-GLUCANOTRANSFERASE (OLIGO-1,4-1,4-GLUCANTRANSFERASE); AMYLO-1,6-GLUCOSIDASE (DEXTRIN 6-ALPHA-D-GLUCOSIDASE)]
9883	22910	36376	1.31	5.0E-01	P35573	SW/ISSPROT	GLYCAGEN DEBRANCHING ENZYME (GLYCAGEN DEBRANCHER) [INCLUDES: 4-ALPHA-GLUCANOTRANSFERASE (OLIGO-1,4-1,4-GLUCANTRANSFERASE); AMYLO-1,6-GLUCOSIDASE (DEXTRIN 6-ALPHA-D-GLUCOSIDASE)]

Page 54 of 546
Table 4

Single Exon Probes Expressed in Bone Marrow

Probe SEQ ID NO:	Exon SEQ ID NO:	ORF SEQ ID NO:	Expression Signal	Most Similar (Top) Hit BLAST E Value	Top Hit Accession No.	Top Hit Database Source	Top Hit Descriptor
10756	23578		1.38	5.0E-01	BE89218.1	EST_HUMAN	601448024F1 NIH_MGC 65 Homo sapiens cDNA clone IMAGE:3849438 5'
12302	25113		6.28	5.0E-01	AF029215.1	NT	Mus musculus MFC OX-2 antigen homolog gene, exons 2-5, and complete cds
13004	25558		2.21	5.0E-01	AL163302.2	NT	Homo sapiens chromosome 21 segment HS21C102
13011	25558		5.6	5.0E-01	O13961	SWISSPROT	NUCLEAR ENVELOPE PROTEIN CUT11
730	13849	26786	2.03	4.9E-01	BF571462.1	EST_HUMAN	602076646F1 NIH_MGC 62 Homo sapiens cDNA clone IMAGE:4243860 5'
1668	14700	27676	2.37	4.9E-01	AJ243955.1	NT	Xenopus laevis mRNA for c-Jun protein, 1978 BP
1921	14945	27941	0.99	4.9E-01	U40869.1	NT	Canis parvulus pulmonary surfactant protein A (SP-a) mRNA, complete cds
5480	18580	31491	1.43	4.9E-01	Q61554	SWISSPROT	FIBRILLIN 1 PRECURSOR
6153	18228	32457	2.67	4.9E-01	AF020331.1	NT	Homo sapiens diacylglycerol kinase 3 (DAGK3) gene, exon 10
6153	18228	32458	2.67	4.9E-01	AF020331.1	NT	Homo sapiens diacylglycerol kinase 3 (DAGK3) gene, exon 10
7685	20643	34007	1.69	4.9E-01	AB040051.1	NT	Oryza sativa subsp. japonica mEF-G mRNA for mitochondrial elongation factor G, complete cds
7972	20911	34301	0.7	4.9E-01	Q10606	SWISSPROT	PUTATIVE UNDECAPRENYL-PHOSPHATE ALPHA-N-ACETYLGLUCOSAMINYLTRANSFERASE
7972	20911	34302	0.7	4.9E-01	Q10606	SWISSPROT	PUTATIVE UNDECAPRENYL-PHOSPHATE ALPHA-N-ACETYLGLUCOSAMINYLTRANSFERASE
9341	22308		1.77	4.9E-01	BF209791.1	EST_HUMAN	601874684F1 NIH_MGC 54 Homo sapiens cDNA clone IMAGE:4102603 5'
9542	22505	35954	0.99	4.9E-01	AW339905.1	EST_HUMAN	hc90c02.x1 Soares_NFL_T_GBC_S1 Homo sapiens cDNA clone IMAGE:2807268 3' similar to TR:095714
9551	26008		2.64	4.9E-01	10946863	NT	O65714 HERC2 ;
10881	23603	37097	0.86	4.9E-01	AF053980.1	NT	Mus musculus unc13 homolog (C. elegans) 1 (Unc13h1), mRNA
10888	23608	37314	0.57	4.9E-01	X60000.1	NT	Mus musculus adenyl cyclase 1 (Adcy1) cDNA, partial cds
12186	25041		1.41	4.9E-01	AF176912.1	NT	H. sapiens DNA for BCL7A gene and BCL7AIGH locus fusion
12997	25963		6.43	4.9E-01	AA613562.1	EST_HUMAN	Homo sapiens neurotrophin-1/B-cell stimulating factor-3 gene, complete cds
5585	18681	31649	8.83	4.8E-01	J02987.1	NT	ng228t11.s1 NCI_CGAP Cor10 Homo sapiens cDNA clone IMAGE:1144652 3'
6836	18889	33184	0.69	4.8E-01	U92882.1	NT	Saccharomyces cerevisiae sporulation protein (SPO11) gene required for meiotic recombination, complete cds
6846	18999		3.82	4.8E-01	AA688978.1	EST_HUMAN	Mus musculus slow skeletal muscle troponin T (Tnni1) gene, complete cds
7538	20501		1.98	4.8E-01	5031650	NT	nu85709.s1 NCI_CGAP_Av1 Homo sapiens cDNA clone IMAGE:1217513
7929	20872	34261	0.78	4.8E-01	AL163209.2	NT	Homo sapiens reproduction 8 (DRS2298E) mRNA
8037	20974	34369	3.66	4.8E-01	AL161492.2	NT	Homo sapiens chromosome 21 segment HS21C009
8037	20974	34370	3.66	4.8E-01	AL161492.2	NT	Arabidopsis thaliana DNA chromosome 4, contig fragment No. 4
8237	21206	34611	0.98	4.8E-01	A1820744.1	EST_HUMAN	Arabidopsis thaliana DNA chromosome 4, contig fragment No. 4
9600	22804		0.97	4.8E-01	BE155148.1	EST_HUMAN	y77f10.y5 Soares breast 2NhtBst Homo sapiens cDNA clone IMAGE:154795 5' similar to contains element
10368	23291		0.56	4.8E-01	BF568633.1	EST_HUMAN	MER6 repetitive element ;
11081	24043		1.75	4.8E-01	X83502.1	NT	PM1-HT0360-201298-004-504 HT0360 Homo sapiens cDNA
							602184267F1 NIH_MGC 42 Homo sapiens cDNA clone IMAGE:4300048 5'
							S.cerevisiae ORFs from chromosome X

Table 4

Single Exon Probes Expressed in Bone Marrow

Probe SEQ ID NO:	Exon SEQ ID NO:	ORF SEQ ID NO:	Expression Signal	Most Similar (Top) Hit BLAST E Value	Top Hit Accession No.	Top Hit Database Source	Top Hit Descriptor
12277	25096		1.65	4.8E-01	AL163227.2	NT	Homo sapiens chromosome 21 segment HS21C027
12503	25737		3.32	4.8E-01	AF227565.1	NT	Trypanosoma cruzi transposon VIP II SIRE repeat region
13088	25786		1.49	4.8E-01	AJ132884.1	NT	Chlamydomonas reinhardtii cop gene, exons 1-8
6864	19721	32986	8.88	4.7E-01	BF217173.1	EST_HUMAN	601883880F1 NIH_MGC_57 Homo sapiens cDNA clone IMAGE:4096387 5'
7241	19976	33273	0.92	4.7E-01	AI204374.1	EST_HUMAN	qf72a09.x1 Soares_testis_NHT Homo sapiens cDNA clone IMAGE:1755544 3'
8197	21167	34577	0.59	4.7E-01	T11414.1	EST_HUMAN	hbc811 Human pancreatic islet Homo sapiens cDNA clone hbc811 5'end
8197	21167	34578	0.59	4.7E-01	T11414.1	EST_HUMAN	hbc811 Human pancreatic islet Homo sapiens cDNA clone hbc811 5'end
9430	22394	35834	0.57	4.7E-01	6981501	NT	Rattus norvegicus Spermin binding protein (Sbp), mRNA
11193	24148		4.76	4.7E-01	AF102873.1	NT	Influenza A virus isolate h151697 hemagglutinin (HA) gene, partial cds
11422	24368	37801	1.78	4.7E-01	U41099.1	NT	Human collagen alpha2(XI) (COL11A2) gene, exons 6 through 16, and partial cds
11613	24551	38111	1.74	4.7E-01	BF528658.1	EST_HUMAN	602043889F1 NCI_OGAP_Brm77 Homo sapiens cDNA clone IMAGE:4181303 5'
11704	24689	38246	1.49	4.7E-01	AW888448.1	EST_HUMAN	RC8-NT0028-240400-011-E08 NT0029 Homo sapiens cDNA
12369	25173		1.83	4.7E-01	BE887763.1	EST_HUMAN	601511333F1 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:3912488 5'
12929	25513		1.38	4.7E-01	BF679515.1	EST_HUMAN	602153926F1 NIH_MGC_83 Homo sapiens cDNA clone IMAGE:4294974 5'
3766	16798	28709	1.53	4.6E-01	BF693300.1	EST_HUMAN	602081103F1 NIH_MGC_81 Homo sapiens cDNA clone IMAGE:4245481 5'
3766	16798	28710	1.83	4.6E-01	BF693300.1	EST_HUMAN	602081103F1 NIH_MGC_81 Homo sapiens cDNA clone IMAGE:4245481 5'
5190	18199		0.93	4.6E-01	M11267.1	NT	Bovine staroid 21-hydroxylase gene (P-450-c21) gene, complete cds
5483	18593	31503	1	4.6E-01	BF313593.1	EST_HUMAN	601900234F1 NIH_MGC_19 Homo sapiens cDNA clone IMAGE:4129472 5'
5483	18593	31504	1	4.6E-01	BF313593.1	EST_HUMAN	601900234F1 NIH_MGC_19 Homo sapiens cDNA clone IMAGE:4129472 5'
5548	18645	31586	3.33	4.6E-01	Q90843	SWISSPROT	INTERFERON REGULATORY FACTOR 3 (IRF-3)
5548	18645	31587	3.33	4.6E-01	Q90843	SWISSPROT	INTERFERON REGULATORY FACTOR 3 (IRF-3)
5825	18721	31880	1.95	4.6E-01	BE734781.1	EST_HUMAN	601688765F1 NIH_MGC_21 Homo sapiens cDNA clone IMAGE:3843637 5'
5839	18735	31897	3.12	4.6E-01	AI247679.1	EST_HUMAN	qf59h02.x1 Soares_fetal_liver_spleen_1NFLS_S1 Homo sapiens cDNA clone IMAGE:1849011 3' similar to
5839	18735	31898	3.12	4.6E-01	AI247679.1	EST_HUMAN	TR-O15338 O15338 BUTYROPHILIN.;
5847	18743	31909	1.46	4.6E-01	P20050	SWISSPROT	qf59h02.x1 Soares_fetal_liver_spleen_1NFLS_S1 Homo sapiens cDNA clone IMAGE:1849011 3' similar to
5730	18824		0.86	4.6E-01	AF212124.1	NT	TR-O15338 O15338 BUTYROPHILIN.;
5821	18911		0.78	4.6E-01	BE817247.1	EST_HUMAN	MEIOSIS SPECIFIC PROTEIN HOP-1
6002	19085	32285	0.51	4.6E-01	D26215.1	NT	Andis schwartzii cytochrome b gene, partial cds; mitochondrial gene for mitochondrial product
							PMO-BN0260-120600-001-F07 BN0260 Homo sapiens cDNA
							Unidentified soil bacteria 16S rRNA gene encoding 16S ribosomal RNA
							Methanobacterium thermoautotrophicum from bases 11657/51 to 1176238 (section 100 of 148) of the
6386	19454	32699	0.92	4.6E-01	AE000894.1	NT	complete genome
6881	19933	33231	0.52	4.6E-01	AF115340.1	NT	Bacillus subtilis Blna (blna) gene, complete cds

Page 56 of 546
Table 4

Single Exon Probes Expressed in Bone Marrow

Probe SEQ ID NO:	Exon SEQ ID NO:	ORF SEQ ID NO:	Expression Signal	Most Similar (Top) Hit BLAST E Value	Top Hit Accession No.	Top Hit Database Source	Top Hit Descriptor
6832	20156	33474	1.43	4.6E-01	U82332.1	NT	Emertcella nidulans NEMPA (nemPA) gene, mitochondrial gene encoding putative mitochondrial protein, complete cds
6832	20156	33475	1.43	4.6E-01	U82332.1	NT	Emertcella nidulans NEMPA (nemPA) gene, mitochondrial gene encoding putative mitochondrial protein, complete cds
7441	25681	33759	0.53	4.6E-01	L07320.1	NT	Murine cytomegalovirus e1 protein gene, complete cds
8001	20940	34333	0.78	4.6E-01	AA493577.1	EST_HUMAN	h04h05.s1 NCL_CGAP_Thy1 Homo sapiens cDNA clone IMAGE:943353 similar to contains Alu repetitive element; contains element L1 repetitive element;
8093	21029		0.53	4.6E-01	AE004031.1	NT	Xylella fastidiosa, section 177 of 229 of the complete genome
8663	21631	35062	13.26	4.6E-01	BF697398.1	EST_HUMAN	60213063F1 NIH_MGC_56 Homo sapiens cDNA clone IMAGE:4287828 5'
9089	22065	35490	0.47	4.6E-01	AA932237.1	EST_HUMAN	0076808.s1 NCL_CGAP_Kid5 Homo sapiens cDNA clone IMAGE:1572087 3' similar to gb:M36341 ADP-RIBOSYLATION FACTOR 4 (HUMAN);
9089	22065	35491	0.47	4.6E-01	AA932237.1	EST_HUMAN	0076808.s1 NCL_CGAP_Kid5 Homo sapiens cDNA clone IMAGE:1572087 3' similar to gb:M36341 ADP-RIBOSYLATION FACTOR 4 (HUMAN);
9656	22569	36048	0.99	4.6E-01	P55202	SWISSPROT	ATRIAL NATRIURETIC PEPTIDE RECEPTOR B PRECURSOR (ANP-B) (ANPRB) (GC-B) (GUANYLATE CYCLASE)
9658	22569	36049	0.99	4.6E-01	P55202	SWISSPROT	ATRIAL NATRIURETIC PEPTIDE RECEPTOR B PRECURSOR (ANP-B) (ANPRB) (GC-B) (GUANYLATE CYCLASE)
10024	22951	36418	0.89	4.6E-01	AF162283.1	NT	Glycine max acetyl-CoA carboxylase (accB-1) gene, complete cds; nuclear gene for chloroplast product
10024	22951	36419	0.89	4.6E-01	AF162283.1	NT	Glycine max acetyl-CoA carboxylase (accB-1) gene, complete cds; nuclear gene for chloroplast product
10335	23259	36736	1.62	4.6E-01	A0195634.1	EST_HUMAN	wg73e12.x1 Soares_NSIF_F8_9W_OT_PA_P_S1 Homo sapiens cDNA clone IMAGE:2370766 3'
10335	23259	36737	1.62	4.6E-01	A0195634.1	EST_HUMAN	wg73e12.x1 Soares_NSIF_F8_9W_OT_PA_P_S1 Homo sapiens cDNA clone IMAGE:2370766 3'
11335	24285		2.28	4.6E-01	P98163	SWISSPROT	PUTATIVE VITELLOGENIN RECEPTOR PRECURSOR (VL)
11343	24283	37818	4.94	4.6E-01	BE185449.1	EST_HUMAN	IL5-HT0730-100500-075-g05 HT0730 Homo sapiens cDNA
11343	24283	37819	4.94	4.6E-01	BE185449.1	EST_HUMAN	IL5-HT0730-100500-075-g05 HT0730 Homo sapiens cDNA
11445	24388	37930	3.88	4.6E-01	BE272325.1	EST_HUMAN	60126068F1 NIH_MGC_9 Homo sapiens cDNA clone IMAGE:2968865 5'
11789	23954	37476	4.41	4.6E-01	AF019369.1	NT	Human thiopurine methyltransferase (TPMT) gene, exon 10 and complete cds
11789	23954	37477	4.41	4.6E-01	AF019369.1	NT	Human thiopurine methyltransferase (TPMT) gene, exon 10 and complete cds
12449	26208		1.69	4.6E-01	D63316.1	EST_HUMAN	HUM105F03B Clontech human fetal brain polyA+ mRNA (#6636) Homo sapiens cDNA clone GEN-105F03 5'
1718	14748		0.92	4.5E-01	BE311420.1	EST_HUMAN	601142105F1 NIH_MGC_14 Homo sapiens cDNA clone IMAGE:3505893 5'
1926	14950	27946	1.34	4.5E-01	AE001931.1	NT	Deinococcus radiodurans R1 section 68 of 229 of the complete chromosome 1
1926	14950	27947	1.34	4.5E-01	AE001931.1	NT	Deinococcus radiodurans R1 section 68 of 229 of the complete chromosome 1

Table 4

Single Exon Probes Expressed in Bone Marrow

Probe SEQ ID NO:	Exon SEQ ID NO:	ORF SEQ ID NO:	Expression Signal	Most Similar (Top) Hit BLAST E Value	Top Hit Accession No.	Top Hit Database Source	Top Hit Descriptor
2881	15040	28857	5.38	4.5E-01	AA677088.1	EST_HUMAN	z55d02.s1 Source_fetal_liver_spleen_1NFLS_S1 Homo sapiens cDNA clone IMAGE:454179 3'
3328	16379	28300	3.85	4.5E-01	Q05793	SWISSPROT	BASEMENT MEMBRANE-SPECIFIC HEPARAN SULFATE PROTEOGLYCAN CORE PROTEIN
3389	16438	28384	1.82	4.5E-01	AF126378.1	NT	PRECURSOR (HSPG) (PERLECAN) (PLG)
4060	17096		1.23	4.5E-01	Q28247	SWISSPROT	Mus musculus DNA polymerase epsilon catalytic subunit (Pole) gene, exons 2 through 12
4101	17135	30030	1.12	4.5E-01	A1708908.1	EST_HUMAN	COLLAGEN ALPHA 5(M) CHAIN
4205	18325		4.08	4.5E-01	AW873495.1	EST_HUMAN	es96a09.x1 Barcode aorta HPLR86 Homo sapiens cDNA clone IMAGE:2353480 3'
4890	17895	30884	1.09	4.5E-01	BE963445.2	EST_HUMAN	h60g02.x1 Soares_NFL_T_GBC_S1 Homo sapiens cDNA clone IMAGE:3041810 3'
5828	18724	31885	1.3	4.5E-01	AW608814.1	EST_HUMAN	901687228R1 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:3868023 3'
6760	19814		1.74	4.5E-01	Q00956	SWISSPROT	QV2-PT0012-140100-031-c09 PT0012 Homo sapiens cDNA
7844	20604	33969	0.83	4.5E-01	M37036.1	NT	COAT PROTEIN
7887	20811	34188	2.84	4.5E-01	A1858949.1	EST_HUMAN	Rat nuclear proteins B23.1 and B23.2
7891	20830	34325	0.51	4.5E-01	P60070	SWISSPROT	w32a02.x1 NCL_OGAP_UH1 Homo sapiens cDNA clone IMAGE:2428618 3' similar to TR-Q82923 Q82923
8650	21618		0.87	4.5E-01	M32861.1	NT	SWISNF COMPLEX 170 KDA SUBUNIT. ;
8746	21714	35137	3.88	4.5E-01	A1848598.1	EST_HUMAN	DNA PRIMASE
							D.melanogaster Shaw2 protein mRNA, complete cds
							tz59g11.x1 NCL_OGAP_O165 Homo sapiens cDNA clone IMAGE:2262844 3'
							POLY-BETA-HYDROXYBUTYRATE POLYMERASE (POLY(3-HYDROXYBUTYRATE) POLYMERASE) (PHB POLYMERASE) (PHB SYNTHASE) (POLY(3-HYDROXYALKANOATE) POLYMERASE) (PHA POLYMERASE) (PHA SYNTHASE) (POLYHYDROXYALKANOIC ACID SYNTHASE)
8905	21871	35287	0.74	4.5E-01	Q52728	SWISSPROT	Homo sapiens hypothetical protein DKFZp547G183 (DKFZp547G183), mRNA
9133	22089		1.72	4.5E-01	11444786	NT	Escherichia coli K-12 MG1655 section 108 of the complete genome
9351	22316	35742	0.78	4.5E-01	AE000218.1	NT	Escherichia coli K-12 MG1655 section 108 of the complete genome
10300	23225		0.89	4.5E-01	9630816	NT	Bombay mol nuclear polyhedrosis virus, complete genome
10861	23781	37281	26.2	4.5E-01	M89006.1	EST_HUMAN	EST02531 Fetal brain, Striatum (cat#836206) Homo sapiens cDNA clone HFBCY17
10861	23781	37282	26.2	4.5E-01	M89006.1	EST_HUMAN	EST02531 Fetal brain, Striatum (cat#836206) Homo sapiens cDNA clone HFBCY17
11212	24165	37695	2.3	4.5E-01	AW591271.1	EST_HUMAN	xx14h01.x1 NCL_OGAP_U3 Homo sapiens cDNA clone IMAGE:2703895 3' similar to SW:INT6_MOUSE
12165	25955		5.3	4.5E-01	BE871461.1	EST_HUMAN	Q84252 VIRAL INTEGRATION SITE PROTEIN INT-6 [1];
12349	25144		1.48	4.5E-01	O18636	SWISSPROT	601449201F1 NIH_MGC_65 Homo sapiens cDNA clone IMAGE:3852361 5'
12452	25211		1.54	4.5E-01	AJ132045.1	NT	OUT AT FIRST PROTEIN
12891	25478		8.22	4.5E-01	11422069	NT	Thelazia annulata sHAT2 gene
2052	15071		2.23	4.4E-01	6880503	NT	Homo sapiens testis-specific kinase 2 (TESK2), mRNA
							Mus musculus integral membrane-associated protein 1 (Itimap1), mRNA
2398	15405	28430	7.02	4.4E-01	P49765	SWISSPROT	VASCULAR ENDOTHELIAL GROWTH FACTOR B PRECURSOR (VEGF-B) (VEGF RELATED FACTOR)

Table 4

Single Exon Probes Expressed in Bone Marrow

Probe SEQ ID NO:	Exon SEQ ID NO:	ORF SEQ ID NO:	Expression Signal	Most Similar (Top) Hit BLAST E Value	Top Hit Accession No.	Top Hit Database Source	Top Hit Descriptor
3326	16377	28298	1.36	4.4E-01	AF058790.1	NT	Rattus norvegicus SynGAP-b mRNA, complete cds
3326	16377	28299	1.36	4.4E-01	AF058790.1	NT	Rattus norvegicus SynGAP-b mRNA, complete cds
3330	16381	28302	1.9	4.4E-01	BF086726.1	EST_HUMAN	791402.Y1 NCI_CGAP_B18 Homo sapiens cDNA clone IMAGE:3393795 5'
4262	17291		1.75	4.4E-01	BE378707.1	EST_HUMAN	601237139F1 NIH_MGC_44 Homo sapiens cDNA clone IMAGE:3609393 5'
5494	18594	31506	1.63	4.4E-01	P04829	SWISSPROT	HISTIDINE-RICH GLYCOPROTEIN PRECURSOR
5494	18594	31506	1.63	4.4E-01	P04829	SWISSPROT	HISTIDINE-RICH GLYCOPROTEIN PRECURSOR
5772	18884	32045	1.77	4.4E-01	S65019.1	NT	much [rats, Sprague-Dawley, sulfur-dioxide-treated tracheal epithelium, mRNA Partial, 390 nt]
5790	18882	32064	1.82	4.4E-01	AV720408.1	EST_HUMAN	AV720408 GLC Homo sapiens cDNA clone GLCCSC12 5'
6064	19145	32356	1.42	4.4E-01	AI198413.1	EST_HUMAN	q622h11.x1 NCI_CGAP_Bm26 Homo sapiens cDNA clone IMAGE:1861125 3' similar to TR:Q29168 Q29169
6064	19145	32357	1.42	4.4E-01	AI198413.1	EST_HUMAN	UNKNOWN PROTEIN ;
6368	19436	32680	1.8	4.4E-01	AW080785.1	EST_HUMAN	q622h11.x1 NCI_CGAP_Cot18 Homo sapiens cDNA clone IMAGE:2585510 3' similar to TR:Q95154 Q95154
6462	19527		1.17	4.4E-01	AA776132.1	EST_HUMAN	AFLATOXIN B1-ALDEHYDE REDUCTASE. ;
7627	20587	33950	0.95	4.4E-01	AE000571.1	NT	ty85d11.s1 Stratiogene schizo brain S11 Homo sapiens cDNA clone IMAGE:970965 3' similar to gb:M16038
8173	21143		12.66	4.4E-01	Z11679.1	NT	Helicobacter pylori 26695 section 49 of 134 of the complete genome
9115	22081	35509	0.74	4.4E-01	AA056427.1	EST_HUMAN	S.tuberosum mRNA for induced stolon tip protein (partial)
9506	22469	35913	0.72	4.4E-01	AF112540.1	NT	z68a03.s1 Stratiogene colon (#337204) Homo sapiens cDNA clone IMAGE:508636 3'
9638	22501	35949	0.56	4.4E-01	AW612578.1	EST_HUMAN	HIV-1 isolate 08107v6 from USA, envelope glycoprotein (env) gene, partial cds
9845	22589	36038	1.21	4.4E-01	O62838	SWISSPROT	h03c08.x1 NCI_CGAP_JQd11 Homo sapiens cDNA clone IMAGE:2954222 5' similar to
10321	23245	36725	2.1	4.4E-01	AI288650.1	EST_HUMAN	SW-MSH6_HUMAN P52701 DNA MISMATCH REPAIR PROTEIN MSH6 ;
10322	23246		3.91	4.4E-01	P28822	SWISSPROT	ZINC FINGER X-CHROMOSOMAL PROTEIN
10457	23379	36872	5.07	4.4E-01	P35590	SWISSPROT	q639f09.x1 NCI_CGAP_Lu6 Homo sapiens cDNA clone IMAGE:1910921 3'
10740	23662	37196	1.27	4.4E-01	S78404.1	NT	GLYCOPROTEIN B PRECURSOR (GLYCOPROTEIN 14)
10740	23662	37157	1.27	4.4E-01	S78404.1	NT	TYROSINE-PROTEIN KINASE RECEPTOR TIE-1 PRECURSOR
12432	25188	31824	5.76	4.4E-01		NT	beta -HKA-H,K-ATPase beta-subunit [rats, Genomic, 8983 nt, segment 2 of 2]
12661	25465	31727	2.83	4.4E-01	6677874	NT	beta -HKA-H,K-ATPase beta-subunit [rats, Genomic, 8983 nt, segment 2 of 2]
12667	25535			4.4E-01	9627742	NT	Mus musculus sodium channel, type X, alpha polypeptide (Scn10a), mRNA
411	13484	26419	2.49	4.3E-01	P54725	SWISSPROT	Autographa californica nucleopolydnavirus, complete genome
411	13484	26420	2.49	4.3E-01	AF155218.1	NT	UV EXCISION REPAIR PROTEIN PROTEIN RAD23 HOMOLOG A (HHR23A)
1607	14639	27616	0.96	4.3E-01	AF155218.1	EST_HUMAN	Callithrix jacchus MW/LW opsin gene, upstream flanking region
					AW866550.1	EST_HUMAN	Callithrix jacchus MW/LW opsin gene, upstream flanking region
							QV4-SN0024-200400-183-b01 SN0024 Homo sapiens cDNA

Table 4

Single Exon Probes Expressed in Bone Marrow

Probe SEQ ID NO:	Exon SEQ ID NO:	ORF SEQ ID NO:	Expression Signal	Most Similar (Top) Hit BLAST E Value	Top Hit Accession No.	Top Hit Database Source	Top Hit Descriptor
2883	15942		1.83	4.3E-01	AW935288.1	EST_HUMAN	CM2-DT0003-010200-077-c01 DT0003 Homo sapiens cDNA
3073	16130	28042	0.94	4.3E-01	AW989477.1	EST_HUMAN	MRO-BN0070-270300-008-g04 BN0070 Homo sapiens cDNA
4178	17209	30095	1.32	4.3E-01	J00308.1	NT	Human somatostatin 1 gene and flanks
4435	13484	28419	1.63	4.3E-01	AF155218.1	NT	Calitritix jecchus MW/LW opsin gene, upstream flanking region
4435	13484	28420	1.63	4.3E-01	AF155218.1	NT	Calitritix jecchus MW/LW opsin gene, upstream flanking region
5154	18164		1.17	4.3E-01	8635250	NT	Xestia c-nigrum granulovirus, complete genome
5291	18236	31157	1.16	4.3E-01	BE780162.1	EST_HUMAN	601468030F1 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:3871255 5'
5438	18540	31449	0.89	4.3E-01	P48634	SWISSPROT	LARGE PROLINE-RICH PROTEIN BAT2 (HLA-B-ASSOCIATED TRANSCRIPT 2)
5438	18540	31450	0.89	4.3E-01	P48634	SWISSPROT	LARGE PROLINE-RICH PROTEIN BAT2 (HLA-B-ASSOCIATED TRANSCRIPT 2)
5990	19075	32273	1.5	4.3E-01	BE181955.1	EST_HUMAN	QV1-HT0638-070500-191-d08 HT0638 Homo sapiens cDNA
6010	19093	32283	1.83	4.3E-01	AF179825.1	NT	Salimiri sclerous olfactory receptor (SSC168) gene, partial cds
6866	19919	33214	4	4.3E-01	AJ001678.1	NT	Coturnix coturnix japonica fnG gene
7049	20071		0.86	4.3E-01	Q33367	SWISSPROT	DNA GYRASE SUBUNIT B
7682	20622		1.83	4.3E-01	BF348001.1	EST_HUMAN	802023134F1 NCI_CGAP_Binf7 Homo sapiens cDNA clone IMAGE:4158298 5'
8116	21053		0.62	4.3E-01	M58843.1	NT	Human lipoprotein associated coagulation inhibitor (LACI) gene, exon 2
8770	21737		2.58	4.3E-01	U97040.1	NT	Methanococcus voltae flagella-related protein C-1 (flaC-fla) genes, complete cds
9610	22814	36068	1.04	4.3E-01	Y14604.1	NT	Erwinia amylovora rcsV gene
10084	23011	36483	1.88	4.3E-01	AW630048.1	EST_HUMAN	ht74et0.y1 NCI_CGAP_GU1 Homo sapiens cDNA clone IMAGE:2868554 5'
10084	23011	36484	1.88	4.3E-01	AW630048.1	EST_HUMAN	ht74et0.y1 NCI_CGAP_GU1 Homo sapiens cDNA clone IMAGE:2868554 5'
10589	23511	37005	0.85	4.3E-01	AW170559.1	EST_HUMAN	xt63s05.x1 Soares_NHCCc_cervical_tumor Homo sapiens cDNA clone IMAGE:2888400 3' similar to
10879	23789	37301	0.48	4.3E-01	H68292.1	EST_HUMAN	TR-000189 O00189 MU-ADAPTIN-RELATED PROTEIN 2 :
11277	20178	33503	1.55	4.3E-01	AF075629.1	NT	yr45s05.s1 Soares fetal liver spleen 1NFSL Homo sapiens cDNA clone IMAGE:208209 3'
11539	24480	38031	1.77	4.3E-01	AW883658.1	EST_HUMAN	Equus caballus microsatellite LEX027
11539	24480	38032	1.77	4.3E-01	AW883658.1	EST_HUMAN	RC3-BN0034-280200-013-c12 BN0034 Homo sapiens cDNA
13055	25592		2.24	4.3E-01	AJ003022.1	NT	RC3-BN0034-280200-013-c12 BN0034 Homo sapiens cDNA
1360	15865	27365	1.77	4.2E-01	Q39102	SWISSPROT	Streptomyces coelicolor whiH gene
3623	16666	29580	4.43	4.2E-01	AE003947.1	NT	CELL DIVISION PROTEIN FISH HOMOLOG PRECURSOR
3651	16684	29609	1.04	4.2E-01	A1280338.1	EST_HUMAN	Xylella fastidiosa, section 93 of 229 of the complete genome
3724	18324		1.22	4.2E-01	N81203.1	EST_HUMAN	qB44501.x1 Soares_NHrMIPu_S1 Homo sapiens cDNA clone IMAGE:1879945 3'
3893	18833	29843	0.69	4.2E-01	AW835527.1	EST_HUMAN	788E1 fetal brain cDNA Homo sapiens cDNA clone 788E1-K similar to R07878, 240498
4007	17046	29853	1.72	4.2E-01	Q04886	SWISSPROT	QV0-L T0015-180200-127-h01 LT0015 Homo sapiens cDNA
4726	17746	30637	4.9	4.2E-01	AA534093.1	EST_HUMAN	SOX-8 PROTEIN
							rt69h01.s1 NCI_CGAP_P710 Homo sapiens cDNA clone IMAGE:887777 similar to gb:MF33600 HLA CLASS II HISTOCOMPATIBILITY ANTIGEN, DR-1 BETA CHAIN (HUMAN);

Table 4

Single Exon Probes Expressed in Bone Marrow

Probe SEQ ID NO:	Exon SEQ ID NO:	ORF SEQ ID NO:	Expression Signal	Most Similar (Top) Hit BLAST E Value	Top Hit Accession No.	Top Hit Database Source	Top Hit Descriptor
4820	17837	30735	3.62	4.2E-01	R13467.1	EST_HUMAN	Y77601.1 Soares infant brain 1N1B Homo sapiens cDNA clone IMAGE:28278 5'
5801	18893	32076	1.38	4.2E-01	BF242055.1	EST_HUMAN	601879721F1 NIH_MGC_55 Homo sapiens cDNA clone IMAGE:4108493 5'
5878	18965	32156	1.23	4.2E-01	AW854182.1	EST_HUMAN	RCS-CT0254-060400-029-g04 CT0254 Homo sapiens cDNA
6329	19398	32841	0.91	4.2E-01	AL163247.2	NT	Homo sapiens chromosome 21 segment HS21C047
7139	20115	33428	9.39	4.2E-01	AU158472.1	EST_HUMAN	AU158472 PLACE2 Homo sapiens cDNA clone PLACE2000470 3'
7139	20115	33429	9.39	4.2E-01	AU158472.1	EST_HUMAN	AU158472 PLACE2 Homo sapiens cDNA clone PLACE2000470 3'
7207	25677	33563	5.51	4.2E-01	S82504.1	NT	Brcal1=breast cancer gene [rats, WF, spleen, Genomic, 419 nt, segment 2 of 2]
7900	20272	33607	5.9	4.2E-01	AL161547.2	NT	Arabidopsis thaliana DNA chromosome 4, contig fragment No. 47
7822	20770	34146	0.52	4.2E-01	AL163252.2	NT	Homo sapiens chromosome 21 segment HS21C062
8325	21294	34708	2.56	4.2E-01	AW957448.1	EST_HUMAN	EST368413 MAGE resequences, MAGE Homo sapiens cDNA
8325	21294	34709	2.56	4.2E-01	AW957448.1	EST_HUMAN	EST368413 MAGE resequences, MAGE Homo sapiens cDNA
8548	21516	34934	0.49	4.2E-01	4758039	NT	Homo sapiens cytochrome c oxidase subunit VIc (COX6C), nuclear gene encoding mitochondrial protein, mRNA
10329	23253		0.72	4.2E-01	AA705007.1	EST_HUMAN	Z85101.s1 Soares fetal liver spleen_1N1FLS_S1 Homo sapiens cDNA clone IMAGE:462849 3'
10541	23463	36958	0.43	4.2E-01	AF181854.1	NT	Lassa virus strain 803213 glycoprotein precursor and nucleoprotein genes, complete cds
10862	23782	37283	1.53	4.2E-01	AW863686.1	EST_HUMAN	MR3-SN0010-280300-103-h07 SN0010 Homo sapiens cDNA
11382	24329	37658	2.39	4.2E-01	AB023489.1	NT	Oryzias latipes OGG7 mRNA for membrane guanylyl cyclase, complete cds
11723	24609	38186	1.77	4.2E-01	BE966485.2	EST_HUMAN	601660352R1 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:3906085 3'
13090	25618		1.4	4.2E-01	AI392837.1	EST_HUMAN	tg10c05.x1 NCI_CGAP_CLL1 Homo sapiens cDNA clone IMAGE:2108360 3'
1098	14140	27060	1.98	4.1E-01	AI905481.1	EST_HUMAN	RC-BT091-210169-142 BT091 Homo sapiens cDNA
1105	14149	27069	1.33	4.1E-01	AV705243.1	EST_HUMAN	AV705243 ADB Homo sapiens cDNA clone ADBAHF08 5'
1105	14149	27100	1.33	4.1E-01	AV705243.1	EST_HUMAN	AV705243 ADB Homo sapiens cDNA clone ADBAHF08 5'
2722	15716	28734	1.43	4.1E-01	7705283	NT	Homo sapiens anaphase-promoting complex subunit 7 (APC7), mRNA
2851	16008	28932	2.07	4.1E-01	AL161538.2	NT	Arabidopsis thaliana DNA chromosome 4, contig fragment No. 38
2951	16008	28933	2.07	4.1E-01	AL161538.2	NT	Arabidopsis thaliana DNA chromosome 4, contig fragment No. 38
3314	16367	29287	0.69	4.1E-01	AA906344.1	EST_HUMAN	q94b08.s1 Soares NFL_T_GBC_S1 Homo sapiens cDNA clone IMAGE:1505943 3'
4303	17332	30212	2.54	4.1E-01	AL249207.1	NT	Rhodococcus sp. AD45 isoG, isoH, isoI, isoJ, isoK, isoL, isoM, isoN, isoO, isoP, isoQ, isoR, isoS, isoT, isoU, isoV, isoW, isoX, isoY, isoZ, and isoF genes
4334	17362		0.84	4.1E-01	AA909287.1	EST_HUMAN	cm33d02.s1 Soares NFL_T_GBC_S1 Homo sapiens cDNA clone IMAGE:1542819 3'
4894	17715	30610	1.71	4.1E-01	AV747880.1	EST_HUMAN	AV747880 NPC Homo sapiens cDNA clone NPCBDF10 5'
6103	19182	32401	4.57	4.1E-01	BF681393.1	EST_HUMAN	602156590F1 NIH_MGC_83 Homo sapiens cDNA clone IMAGE:4287319 5'
6877	19930	33227	0.68	4.1E-01	U02238.1	NT	Mus musculus NIH 3T3 chemokine ratites (Soyes) gene, complete cds
7668	20625	33989	2.98	4.1E-01	U67535.1	NT	Methanococcus jannaschii section 77 of 150 of the complete genome
8370	21339	34750	1.16	4.1E-01	BF574804.1	EST_HUMAN	602133261F1 NIH_MGC_81 Homo sapiens cDNA clone IMAGE:4288238 5'

Page 61 of 546
Table 4

Single Exon Probes Expressed in Bone Marrow

Probe SEQ ID NO:	Exon SEQ ID NO:	ORF SEQ ID NO:	Expression Signal	Most Similar (Top) Hit BLAST E Value	Top Hit Accession No.	Top Hit Database Source	Top Hit Descriptor
9445	22409	35848	1.45	4.1E-01	6755521	NT	Mus musculus signalling intermediate in Toll pathway-evolutionarily conserved (Slpco-pending), mRNA
9923	22807		0.58	4.1E-01	AF160597.1	NT	Vodveto gymnocaustus Vgym560 cytochrome b (cytb) gene, complete cds; mitochondrial gene for mitochondrial product
10628	23548		1.18	4.1E-01	AL139076.2	NT	Campylobacter jejuni NCTC11168 complete genome, segment 3/6
10775	23696	37194	1.1	4.1E-01	AV649579.1	EST_HUMAN	AV649579 GLC Homo sapiens cDNA clone GLC8VD12 3'
10873	23793	37294	0.51	4.1E-01	P18584	SWISSPROT	PROBABLE SERINE PROTEASE DO-LIKE PRECURSOR (59 KDA IMMUNOGENIC PROTEIN) (SK59)
10873	23793	37295	0.51	4.1E-01	P18584	SWISSPROT	PROBABLE SERINE PROTEASE DO-LIKE PRECURSOR (59 KDA IMMUNOGENIC PROTEIN) (SK59)
10943	23883		2.1	4.1E-01	BF346382.1	EST_HUMAN	CM2-HT0137-200889-010-g08 HT0137 Homo sapiens cDNA
11188	24144	37677	39.55	4.1E-01	X58700.1	NT	Zee mays ZMIPMS2 gene for 19 kDa zahn protein
12751	25931		3	4.1E-01	D87675.1	NT	Homo sapiens DNA for amyloid precursor protein, complete cds
142	15633		0.65	4.0E-01	AW847123.1	EST_HUMAN	RC2-CT0201-290899-012-d10 CT0201 Homo sapiens cDNA
1040	14085	27036	0.71	4.0E-01	8404656	NT	Lequeus rubellus mitochondrion, complete genome
1342	14377	27346	1.17	4.0E-01	AF203478.1	NT	Drosophila melanogaster Dalmation (dmt) mRNA, complete cds
1481	14514		5.11	4.0E-01	6678258	NT	Mus musculus platelet derived growth factor receptor, beta polypeptide (Pdgfrb), mRNA
2022	15883	28054	1.12	4.0E-01	Z96933.1	NT	Ascorbolus immerus mesc2 gene
2022	15883	28055	1.12	4.0E-01	Z96933.1	NT	Ascorbolus immerus mesc2 gene
2167	16183	28203	1.09	4.0E-01	AE001831.1	NT	Deinococcus radiodurans R1 section 68 of 229 of the complete chromosome 1
2167	16183	28204	1.09	4.0E-01	AE001831.1	NT	Deinococcus radiodurans R1 section 68 of 229 of the complete chromosome 1
2816	13246	26176	1.27	4.0E-01	6678490	NT	Mus musculus ubiquitin-protein ligase e3 componen n-recogin (Ubr1), mRNA
2979	16037	28969	1.34	4.0E-01	AL163280.2	NT	Homo sapiens chromosome 21 segment HS21C080
2979	16037	28960	1.34	4.0E-01	AL163280.2	NT	Homo sapiens chromosome 21 segment HS21C080
							Streptococcus pneumoniae Y1C (y1C), Y1D (y1D), penicillin-binding protein 2x (pbp2x), and undecaprenyl- phosphate-UDP-MurNAc-pentapeptide phospho-MurNAc-pentapeptide transferase (mraY) genes, complete cds
3709	16752	29668	2.32	4.0E-01	AF068903.1	NT	Ovis aries partial JD2 gene for T cell receptor delta chain (TCRDJ2), exon 1
3840	16880	29783	3.28	4.0E-01	AJ277511.1	NT	Ovis aries partial JD2 gene for T cell receptor delta chain (TCRDJ2), exon 1
3840	16880	29784	3.28	4.0E-01	AJ277511.1	NT	Ovis aries partial JD2 gene for T cell receptor delta chain (TCRDJ2), exon 1
4855	17872		9.38	4.0E-01	Q31849	SWISSPROT	NADH-PLASTOQUINONE OXIDOREDUCTASE CHAIN 5, CHLOROPLAST
6015	19098	32299	1.14	4.0E-01	AW970610.1	EST_HUMAN	EST382691 MAGE resequences, MAGK Homo sapiens cDNA
6228	19300	32533	0.51	4.0E-01	BF243741.1	EST_HUMAN	601877853F1 NIH_MGC_55 Homo sapiens cDNA clone IMAGE:4108221 5'
							STRUCTURAL POLYPEPTIDE (P-130) [CONTAINS: COAT PROTEIN C; SPIKE GLYCOPROTEINS E3, E2 AND E1; 8 KD PEPTIDE]
6578	19638	32804	0.92	4.0E-01	P27285	SWISSPROT	Homo sapiens OCTN2 gene, complete cds
8345	21314	34729	0.8	4.0E-01	AB016825.1	NT	Homo sapiens OCTN2 gene, complete cds
9360	22325	35753	0.98	4.0E-01	AA323289.1	EST_HUMAN	EST26068 Cerebellum II Homo sapiens cDNA 5' end similar to EST containing Alu repeat

Page 62 of 546
Table 4
Single Exon Probes Expressed in Bone Marrow

Probe SEQ ID NO:	Exon SEQ ID NO:	ORF SEQ ID NO:	Expression Signal	Most Similar (Top) Hit BLAST E Value	Top Hit Accession No.	Top Hit Database Source	Top Hit Descriptor
11885	24776		2	4.0E-01	BF030282.1	EST_HUMAN	601558283F1 NIH_MGC_58 Homo sapiens cDNA clone IMAGE:3828082 5'
12030	24906		1.75	4.0E-01	L76080.1	NT	Synechocystis sp. PCC 9413 transposase gene, complete cds
12448	25783		3.03	4.0E-01	AL163300.2	NT	Homo sapiens chromosome 21 segment HS21C100
12968	25636		1.8	4.0E-01	P36049	SWISSPROT	HYPOTHETICAL 49.7 KD PROTEIN IN GIN2-STE3 INTERGENIC REGION
1378	14412	27383	1.85	3.9E-01	AF206618.1	NT	Gorilla gorilla carboxyl-ester lipase (CEL) gene, complete cds
2650	15947	28670	3.82	3.9E-01	AB033019.1	NT	Homo sapiens mRNA for KIAA1193 protein, partial cds
2716	15710	28726	4.26	3.9E-01	X82032.1	NT	H. sapiens B-myb gene
2718	15710	28727	4.26	3.9E-01	X82032.1	NT	H. sapiens B-myb gene
3113	16170	29080	4.56	3.9E-01	AJ225896.1	NT	Sinorhizobium meliloti egf, syrB2, cys3 genes and ori3
4106	17140	30035	1.25	3.9E-01	BF582011.1	EST_HUMAN	761401.x1 NCI_CGAP_Br16 Homo sapiens cDNA clone IMAGE:3339188 3'
5027	18041	30824	1.74	3.9E-01	BE728687.1	EST_HUMAN	6015533948F1 NIH_MGC_20 Homo sapiens cDNA clone IMAGE:3833689 5'
6042	19124	32328	5.95	3.9E-01	BF208036.1	EST_HUMAN	601862362F1 NIH_MGC_53 Homo sapiens cDNA clone IMAGE:4082055 5'
							Homo sapiens zinc finger protein 82 (ZFP82), expressed-Xq28STS protein (XQ28ORF), and biglycan (BGN) genes, complete cds; and plasma membrane calcium ATPase isoform 3 (PMCA3) gene, partial cds
8411	19479	32728	0.86	3.9E-01	U82695.2	NT	Homo sapiens propro dipeptidyl peptidase I (DPP-I) gene, complete cds
8286	21255	34684	0.8	3.9E-01	U78415.1	NT	CM3-CT0105-170899-004-b08 CT0105 Homo sapiens cDNA
8213	22179	35610	0.73	3.9E-01	AW177011.1	EST_HUMAN	602019944F1 NCI_CGAP_Brm67 Homo sapiens cDNA clone IMAGE:4155322 5'
8222	22188		0.88	3.9E-01	BF348634.1	EST_HUMAN	xn86cd04.x1 Soares_NFL_T_GBC_S1 Homo sapiens cDNA clone IMAGE:2701351 3' similar to TR:094821
9590	22552	36003	1.41	3.9E-01	AW195888.1	EST_HUMAN	O84821 KIAA0713 PROTEIN ;
							wp76a02.x1 NCI_CGAP_Brm25 Homo sapiens cDNA clone IMAGE:2467658 3' similar to
9803	22855	36316	1.83	3.9E-01	A1837337.1	EST_HUMAN	SW:RFX5_HUMAN P48382 BINDING REGULATORY FACTOR ;
10237	23162	36650	2.88	3.9E-01	M19879.1	NT	Human clabindin 27 gene, exons 10 and 11, and L1 and Alu repeats
10305	23250		0.46	3.9E-01	11465620	NT	Porphyra purpurea mitochondrion, complete genome
10527	23449	36947	0.82	3.9E-01	D86722.1	NT	Nicotiana tabacum mRNA for TATA binding protein (TBP), complete cds
10865	23885	37397	0.48	3.9E-01	M18440.1	NT	Human beta-B2-crystallin (B2-1) gene, exon 4, partial cds
12219	25868		4.08	3.9E-01	AF304354.1	NT	Homo sapiens proteoglycan 3 (PRG3) gene, complete cds
12344	25142		2.01	3.9E-01	Q61670	SWISSPROT	HOMEOBOX PROTEIN HLX1
12840	25452		1.49	3.9E-01	11433335	NT	Homo sapiens hypothetical protein FLJ10583 (FLJ10583), mRNA
161	13264		9.02	3.8E-01	7019488	NT	Homo sapiens protein kinase PKNbeta (pkribeta), mRNA
506	13577		0.8	3.8E-01	AB028291.1	NT	Mus musculus pom-1 mRNA for pericentriolar material-1, complete cds
1886	14911		1.19	3.8E-01	AE003870.1	NT	Xylella fastidiosa, section 16 of 229 of the complete genome
2465	15469	28493	1.24	3.8E-01	U41846.1	NT	Ceanothus brittanii trigonate acetylcholinesterase (ace-1) gene, complete cds
2577	15578	28597	2.26	3.8E-01	AF214117.1	NT	Arabidopsis thaliana putative c-myb-like transcription factor (MYB3R-3) mRNA, complete cds

Table 4

Single Exon Probes Expressed in Bone Marrow

Probe SEQ ID NO:	Exon SEQ ID NO:	ORF SEQ ID NO:	Expression Signal	Most Similar (Top) Hit BLAST E Value	Top Hit Accession No.	Top Hit Database Source	Top Hit Descriptor
2840	15900	28962	4.89	3.8E-01	6078002	NT	Mus musculus solute carrier family 1, member 6 (Slc1a6), mRNA
3016	16073		0.88	3.8E-01	AJ251057.1	NT	Human immunodeficiency virus type 1 complete genome (isolate 98SE-MIP1213)
3081	16118	29033	2.11	3.8E-01	AF043383.1	NT	Pleurococcus americanus aminopeptidase N (ampN) gene, partial cds
3495	16542	29468	9.24	3.8E-01	AL161518.2	NT	Arabidopsis thaliana DNA chromosome 4, contig fragment No. 30
3549	16595		0.75	3.8E-01	AL161518.2	EST_HUMAN	w38b12.x1 Scores_NFL_T_GBC_S1 Homo sapiens cDNA clone IMAGE:2357855 3'
3568	16595		0.87	3.8E-01	AB07219.1	EST_HUMAN	w38b12.x1 Scores_NFL_T_GBC_S1 Homo sapiens cDNA clone IMAGE:2357855 3'
3769	16811	29720	0.99	3.8E-01	BE154080.1	EST_HUMAN	PMO-HT0339-200400-010-G01 HT0339 Homo sapiens cDNA
6691	18788	31957	1.08	3.8E-01	Q04888	SWISSPROT	TRANSCRIPTION FACTOR SOX-10
6473	19338		0.6	3.8E-01	S46825.1	NT	prion protein [mink, Genomic, 2448 nt]
6779	19834	33117	5.49	3.8E-01	BE072398.1	EST_HUMAN	QV3-BT0537-271299-049-602 BT0537 Homo sapiens cDNA
6924	20148	33468	4.78	3.8E-01	AJ374601.1	EST_HUMAN	ts54f11.x1 Scores_total_fetus_Nb2HF8_9w Homo sapiens cDNA clone IMAGE:2047917 3' similar to contains Alu repetitive element
7126	20059	33365	1.24	3.8E-01	AL161513.2	NT	Arabidopsis thaliana DNA chromosome 4, contig fragment No. 25
7743	20697	34063	0.51	3.8E-01	AA626274.1	EST_HUMAN	z88c05.s1 Scores_testis_NHT Homo sapiens cDNA clone IMAGE:745084 3'
7759	20712		4.27	3.8E-01	X61597.1	NT	M.musculus gene for kallikrein-binding protein
7988	20837	34331	0.49	3.8E-01	V00883.1	NT	Yeast mitochondrial gene for ATPase (genes cit-2 and cit-4)
8640	21608	35031	0.45	3.8E-01	M81385.1	NT	Mouse liver receptor homologous protein (LRH-1) mRNA, complete cds
8903	21869	35294	2.04	3.8E-01	AB046851.1	NT	Homo sapiens mRNA for KIAA1631 protein, partial cds
8972	21938	35363	0.78	3.8E-01	11441284	NT	Homo sapiens FOS-like antigen-1 (FOSL1), mRNA
9169	22135	35561	1.47	3.8E-01	AL163278.2	NT	Homo sapiens chromosome 21 segment HS21C079
9819	22740		6.02	3.8E-01	T06419.1	EST_HUMAN	ye33h08.r1 Scores fetal liver spleen 1NFLS Homo sapiens cDNA clone IMAGE:120539 5' similar to contains Alu repetitive element; contains P1R5 repetitive element
11862	24744		3.41	3.8E-01	BE718219.1	EST_HUMAN	RC0-HT0841-040800-032-b12 HT0841 Homo sapiens cDNA
12000	24877	38473	2.61	3.8E-01	R42550.1	EST_HUMAN	yf92h11.s1 Scores infant brain 1NIB Homo sapiens cDNA clone IMAGE:30289 3'
12000	24877	38474	2.61	3.8E-01	R42550.1	EST_HUMAN	yf92h11.s1 Scores infant brain 1NIB Homo sapiens cDNA clone IMAGE:30289 3'
12433	25199		4.23	3.8E-01	AE001124.1	NT	Borrelia burgdorferi (section 10 of 70) of the complete genome
12557	25888		2.63	3.8E-01	U94788.1	NT	Human p53 (TP53) gene, complete cds
12667	25343		2.84	3.8E-01	BE820268.1	EST_HUMAN	QV3-ET0063-190700-271-605 ET0063 Homo sapiens cDNA
13056	25884		2.03	3.8E-01	AF291483.1	NT	y642b11.s1 Strabagene fetal spleen (#837205) Homo sapiens cDNA clone IMAGE:73821 3' similar to similar to gb:AD6977 SERUM ALBUMIN PRECURSOR (HUMAN)
13063	25905	31422	1.4	3.8E-01	T54787.1	EST_HUMAN	Mus musculus vomeronasal receptor V1RA4 (V1ra4) gene, complete cds
13080	25611	31690	1.57	3.8E-01	AF194972.1	NT	y642b11.s1 Strabagene fetal spleen (#837205) Homo sapiens cDNA clone IMAGE:73821 3' similar to similar to gb:AD6977 SERUM ALBUMIN PRECURSOR (HUMAN)
2490	15493	28518	12.56	3.7E-01	AB037831.1	NT	Mus musculus developmental control protein mRNA, partial cds
3474	16520	29442	11.71	3.7E-01	AF056336.1	NT	Homo sapiens mRNA for KIAA1410 protein, partial cds
							Danio rerio bone morphogenetic protein 4 precursor (BMP4) gene, complete cds

Table 4

Single Exon Probes Expressed in Bone Marrow

Probe SEQ ID NO:	Exon SEQ ID NO:	ORF SEQ ID NO:	Expression Signal	Most Similar (Top) Hit BLAST E Value	Top Hit Accession No.	Top Hit Database Source	Top Hit Descriptor
3884	16924	28833	0.7	3.7E-01	AA316482.1	EST_HUMAN	EST21715 Adrenal gland tumor Homo sapiens cDNA 5' end
4257	17286	30188	0.92	3.7E-01	AI218707.1	EST_HUMAN	ak38c07.x1 Soares NSF_F8_9W_OT_PA_P_S1 Homo sapiens cDNA clone IMAGE:1510188 3'
4348	17375	30255	1.84	3.7E-01	AW678037.1	EST_HUMAN	MF3-O10007-080300-104-b02 OT0007 Homo sapiens cDNA
4416	17443	30334	3	3.7E-01	AE002408.1	NT	Neisseria meningitidis serogroup B strain MC58 section 50 of 208 of the complete genome
5857	18947	32132	1.16	3.7E-01	AF135187.1	NT	Homo sapiens interferon-induced protein p78 (MX1) gene, complete cds
6060	19141	32353	1.35	3.7E-01	AL163278.2	NT	Homo sapiens chromosome 21 segment HS21G078
6859	19716	32883	0.66	3.7E-01	M10808.1	NT	Chicken (White leghorn) delta-1 and delta-2 crystallin genes, complete cds
6880	19737		0.77	3.7E-01	L10363.1	NT	Mus saxicola hemoglobin mRNA, complete cds
7350	20320	33687	3.48	3.7E-01	11525843	NT	Homo sapiens tumor endothelial marker 7 precursor (TEM7), mRNA
7658	20618	33983	0.65	3.7E-01	BE873743.1	EST_HUMAN	6014838877.1 NIH_MGC_69 Homo sapiens cDNA clone IMAGE:3888652 5'
7658	20618	33984	0.65	3.7E-01	BE873743.1	EST_HUMAN	6014838877.1 NIH_MGC_69 Homo sapiens cDNA clone IMAGE:3888652 5'
8059	21008	34404	0.71	3.7E-01	T68802.1	EST_HUMAN	yes0a07.r3 Soares fetal liver spleen 1NFS Homo sapiens cDNA clone IMAGE:38324 5'
8872	21840	35064	1.93	3.7E-01	11436739	NT	Homo sapiens chromosome 12 open reading frame 4 (C12ORF4), mRNA
8872	21840	35065	1.93	3.7E-01	11436739	NT	Homo sapiens chromosome 12 open reading frame 4 (C12ORF4), mRNA
8708	21676	35101	0.69	3.7E-01	AA902812.1	EST_HUMAN	ak33b11.s1 NCI_CGAP_Lu5 Homo sapiens cDNA clone IMAGE:1516701 3'
8558	22518		3.78	3.7E-01	AL271388.1	NT	Gallus gallus mRNA for beta-cardiac 15,15'-dioxygenase (bcDO gene)
10530	23452		0.52	3.7E-01	K00691.1	NT	mouse Ig germline alpha membrane exon region
10570	23492	36994	3.65	3.7E-01	A1336411.1	EST_HUMAN	q46b07.x1 Soares fetal lung_Nb-L19W Homo sapiens cDNA clone IMAGE:1950997 3'
11205	24159	37689	1.9	3.7E-01	X05668.1	NT	Rabbit mRNA for fast skeletal muscle myosin heavy chain (MHC)
11569	24316	37842	2.81	3.7E-01	A1297357.1	NT	Homo sapiens partial LIMD1 gene for LIM domains containing protein 1 and KIAA0851 gene
11369	24316	37843	2.81	3.7E-01	A1297357.1	NT	Homo sapiens partial LIMD1 gene for LIM domains containing protein 1 and KIAA0851 gene
11784	23949	37470	2.34	3.7E-01	X04122.1	NT	Bovine mRNA for terminal deoxynucleotidyltransferase (TdT) (EC 2.7.7.31)
12014	24891		1.53	3.7E-01	AA973540.1	EST_HUMAN	cc46d03.e1 NCI_CGAP_Lu5 Homo sapiens cDNA clone IMAGE:1569221 3' similar to gb1M77698
12060	24933		3.22	3.7E-01	6877678	NT	TRANSCRIPTIONAL REPRESSOR PROTEIN YY1 (HUMAN);
12136	25501		2.6	3.7E-01	J04982.1	NT	Mus musculus retinoblastoma 1 (Rb1), mRNA
12308	25117		4.23	3.7E-01	AJ243525.1	NT	Human heart/skeletal muscle ATP/ADP translocase (ANT1) gene, complete cds
12764	25406		1.86	3.7E-01	AL121154.1	EST_HUMAN	Chlamydomonas reinhardtii partial omp1 gene for outer membrane protein 1
12828	25447	31722	2.71	3.7E-01	Y18000.1	NT	DKFZp762K075.J1 762 (synonym: Inmel2) Homo sapiens cDNA clone DKFZp762K075 5'
987	14048		11.36	3.6E-01	U69241.1	NT	Homo sapiens NF-2 gene
1317	14352	27320	2.66	3.6E-01	T80255.1	EST_HUMAN	Human mlb2 gene, partial cds
1317	14352	27321	2.66	3.6E-01	T80255.1	EST_HUMAN	Yd03e05.r1 Soares infant brain 1N1B Homo sapiens cDNA clone IMAGE:24443 5'
1931	14955	27851	6.09	3.6E-01	AW590184.1	EST_HUMAN	Yd03e05.r1 Soares infant brain 1N1B Homo sapiens cDNA clone IMAGE:24443 5'
1931	14955	27852	6.09	3.6E-01	AW590184.1	EST_HUMAN	hg33f02.x1 NCI_CGAP_GC8 Homo sapiens cDNA clone IMAGE:2947419 3'

Table 4

Single Exon Probes Expressed in Bone Marrow

Probe SEQ ID NO:	Exon SEQ ID NO:	ORF SEQ ID NO:	Expression Signal	Most Similar (Top) Hit BLAST E Value	Top Hit Accession No.	Top Hit Database Source	Top Hit Descriptor
1988	14989	27892	6.4	3.6E-01	AF216207.1	NT	Mus musculus ribosomal protein S19 (Rps19) gene, complete cds
2399	15408		3.58	3.6E-01	X76725.1	NT	P. irregularis (P3804) gene for actin
2483	15487	28510	1.29	3.6E-01	L05435.1	NT	Rattus norvegicus synaptic vesicle protein (SV2) mRNA, complete cds
2483	15487	28511	1.29	3.6E-01	L05435.1	NT	Rattus norvegicus synaptic vesicle protein (SV2) mRNA, complete cds
2495	15498	28524	1.63	3.6E-01	AW812033.1	EST_HUMAN	RC5-ST0171-181089-011-g07 ST0171 Homo sapiens cDNA
							PROTEIN-L-ISOASPARTATE O-METHYLTRANSFERASE (PROTEIN-BETA-ASPARTATE METHYLTRANSFERASE) (PIMT) (PROTEIN L-ISOASPARTYL METHYLTRANSFERASE) (L-ISOASPARTYL PROTEIN CARBOXYL METHYLTRANSFERASE)
2638	16637	28680	1.23	3.6E-01	P24206	SWISSPROT	Drosophila melanogaster sugar transporter 3 (suc3) mRNA, complete cds
2910	18322		5.89	3.6E-01	AF189485.1	NT	H. sapiens serotonin transporter gene, exons 9 and 10
3483	18529	29453	2.01	3.6E-01	X76758.1	NT	H. sapiens serotonin transporter gene, exons 9 and 10
3483	18529	29454	2.01	3.6E-01	X76758.1	NT	H. sapiens serotonin transporter gene, exons 9 and 10
4436	17462	30351	1.14	3.6E-01	BE707883.1	EST_HUMAN	RC1-HT0545-150600-014-b712 HT0545 Homo sapiens cDNA
4767	17805	30697	0.72	3.6E-01	Y11526.1	NT	Zmays mRNA for casein kinase II alpha subunit
5045	18058	30937	2.61	3.6E-01	AW339393.1	EST_HUMAN	hs02g04.x1 NCI_CGAP_Lu24 Homo sapiens cDNA clone IMAGE:2872568 3'
5136	18145	31025	0.66	3.6E-01	BE067689.1	EST_HUMAN	MR4-BT0358-270300-005-G10 BT0358 Homo sapiens cDNA
5274	18280	31143	0.91	3.6E-01	AF067959.1	NT	Gallus gallus homeodomain protein HOXD-3 mRNA, complete cds
5274	18280	31144	0.91	3.6E-01	AF067959.1	NT	Gallus gallus homeodomain protein HOXD-3 mRNA, complete cds
5458	18558	31469	0.71	3.6E-01	AJ008665.1	NT	Homo sapiens lipo gene intron 5
							FORMATE HYDROGENLYASE SUBUNIT 5 PRECURSOR (FHL SUBUNIT 5) (HYDROGENASE-3 COMPONENT E)
6205	19279	32512	1.1	3.6E-01	P16431	SWISSPROT	Homo sapiens PHF5 gene
6823	19681	32958	1.8	3.6E-01	Y10186.1	NT	Y74408.1 Soares fetal liver spleen 1NFLS Homo sapiens cDNA clone IMAGE:275987 5'
7355	20325		3.63	3.6E-01	R94080.1	EST_HUMAN	w72c10.x1 Soares_thymus_NH5Th Homo sapiens cDNA clone IMAGE:2513010 3' similar to TR:O15117
							O15117 FYN BINDING PROTEIN. [1]:
7500	20465	33828	1.88	3.6E-01	AW027174.1	EST_HUMAN	SCO-SPONDIN
8568	21534	34954	0.64	3.6E-01	P36167	SWISSPROT	Arabidopsis thaliana DNA chromosome 4, config fragment No. 79
8622	21590	35008	14.05	3.6E-01	AL161583.2	NT	
							Human hereditary haemochromatosis region, histone 2A-like protein gene, hereditary haemochromatosis (HLA-H) gene, RefSeq gene, and sodium phosphate transporter (NPT3) gene, complete cds
9330	22285	35724	0.53	3.6E-01	U91328.1	NT	
							Human hereditary haemochromatosis region, histone 2A-like protein gene, hereditary haemochromatosis (HLA-H) gene, RefSeq gene, and sodium phosphate transporter (NPT3) gene, complete cds
9330	22285	35725	0.53	3.6E-01	U91328.1	NT	
							Homo sapiens lysosomal-associated membrane protein 2 (LAMP2), transcript variant LAMP2A, mRNA
9855	22320	35746	2.94	3.6E-01	4504956	NT	

Table 4

Single Exon Probes Expressed in Bone Marrow

Probe SEQ ID NO:	Exon SEQ ID NO:	ORF SEQ ID NO:	Expression Signal	Most Similar (Top) Hit BLAST E Value	Top Hit Accession No.	Top Hit Database Source	Top Hit Descriptor
9355	22320	35747	2.84	3.6E-01	4504956	NT	Homo sapiens lysosomal-associated membrane protein 2 (LAMP2), transcript variant LAMP2A, mRNA
9647	22610	35959	1.12	3.6E-01	AL163204.2	NT	Homo sapiens chromosome 21 segment HS21C004
9754	22695	36151	0.93	3.6E-01	X17550.1	NT	D. melanogaster singed gene, exons 3, 4, 5 & 6
9754	22695	36152	0.93	3.6E-01	X17550.1	NT	D. melanogaster singed gene, exons 3, 4, 5 & 6
9824	22673		0.54	3.6E-01	X62825.1	NT	C. perfingens plc gene for phosphatase C upstream region containing bent DNA fragment
10222	23147	36838	10.66	3.6E-01	Q53194	SWISSPROT	PROBABLE PEPTIDE ABC TRANSPORTER ATP-BINDING PROTEIN Y4TS
10352	23278	36750	0.48	3.6E-01	AW752901.1	EST_HUMAN	MR2-CT0222-211099-002-b10 CT0222 Homo sapiens cDNA
10352	23278	36751	0.48	3.6E-01	AW752901.1	EST_HUMAN	MR2-CT0222-211099-002-b10 CT0222 Homo sapiens cDNA
11282	24242	37769	3.04	3.6E-01	BE902390.1	EST_HUMAN	601678418F1 NIH_MGC_21 Homo sapiens cDNA clone IMAGE:3958997 5'
11442	24385	37925	3.26	3.6E-01	AB004293.1	NT	Arabidopsis thaliana mRNA for SigB, complete cds
11772	23927	37448	5.83	3.6E-01	AE000856.1	NT	Methanobacterium thermoautotrophicum from bases 702375 to 714311 (section 62 of 148) of the complete genome
12174	25989		2.05	3.6E-01	Y19210.1	NT	Homo sapiens hHb5 gene for hair keratin, exons 1 to 9
12247	25078		1.42	3.6E-01	D90901.1	NT	Synechocystis sp. PCC6803 complete genome, 3/27, 271600-402289
12257	25093		8.35	3.6E-01	AE000335.1	NT	Escherichia coli K-12 MG1655 section 225 of 400 of the complete genome
12417	25187		6.26	3.6E-01	U68888.1	NT	Mus musculus Enr1 mRNA, complete cds
12770	25410		1.97	3.6E-01	11432568	NT	Homo sapiens myeloid/lymphoid or mixed-lineage leukemia (t(11q23) (Drosophila) homolog), translocated to, 10 (AF10), mRNA
13033	25934		3.33	3.6E-01	AW190229.1	EST_HUMAN	x80e11.x1 NCI_CGAP_Pan1 Homo sapiens cDNA clone IMAGE:2679116 3' similar to gb:K00568 TUBULIN ALPHA-1 CHAIN (HUMAN);
115	13226	26160	1.42	3.6E-01	AL161536.2	NT	Arabidopsis thaliana DNA chromosome 4, contig fragment No. 36
210	13311	26238	3.29	3.6E-01	6878933	NT	Mus musculus mannose receptor, C type 2 (Mrc2), mRNA
726	13787	26722	1.48	3.6E-01	7708136	NT	Homo sapiens GAP-like protein (LOC51306), mRNA
726	13787	26723	1.48	3.6E-01	7708136	NT	Homo sapiens GAP-like protein (LOC51306), mRNA
780	13940	26785	4.95	3.6E-01	BF128798.1	EST_HUMAN	601811080R1 NIH_MGC_48 Homo sapiens cDNA clone IMAGE:4053951 3'
1623	14656	27634	0.96	3.6E-01	BF103688.1	EST_HUMAN	601894653F2 NIH_MGC_19 Homo sapiens cDNA clone IMAGE:4124244 5'
1646	14678	27651	2.57	3.6E-01	U35776.1	NT	Rattus norvegicus ADP-ribosylation factor-directed GTPase activating protein mRNA, complete cds
2281	15303	28328	1.28	3.6E-01	P06788	SWISSPROT	HOMEOBOX PROTEIN HOX-A4 (HOX-1.4) (H1-3)
2613	15899	28636	1.13	3.6E-01	AA223252.1	EST_HUMAN	zr08a09.s1 Striatagene NT2 neuronal precursor 937230 Homo sapiens cDNA clone IMAGE:850872 3'
3318	16858		7.27	3.6E-01	AA642138.1	EST_HUMAN	m60403.s1 NCI_CGAP_Lym3 Homo sapiens cDNA clone IMAGE:1172357 3'
4291	17320	30199	2.18	3.6E-01	AF071253.1	NT	Danio rerio homeobox protein (hox5b) gene, complete cds

Table 4

Single Exon Probes Expressed in Bone Marrow

Probe SEQ ID NO:	Exon SEQ ID NO:	ORF SEQ ID NO:	Expression Signal	Most Similar (Top) Hit BLAST E Value	Top Hit Accession No.	Top Hit Database Source	Top Hit Descriptor
4922	17939	30831	0.7	3.5E-01	N81203.1	EST_HUMAN	788IE1 fetal brain cDNA Homo sapiens cDNA clone 788IE1-K similar to R07878, Z40498
4989	17984	30874	8.46	3.5E-01	M18349.1	NT	Rat leukocyte common antigen (L-CA) gene, exons 1 through 5
5407	18510	31387	0.84	3.5E-01	Q96887	SWISSPROT	EARLY E2A DNA-BINDING PROTEIN
5407	18510	31388	0.84	3.5E-01	Q96887	SWISSPROT	EARLY E2A DNA-BINDING PROTEIN
5628	18726	31886	1.36	3.5E-01	D42045.1	NT	Human mRNA for KIAA0086 gene, complete cds
6365	19434		0.77	3.5E-01	AW863916.1	EST_HUMAN	PM4-SN0012-030400-001-att1 SN0012 Homo sapiens cDNA
6548	19609	32871	0.62	3.5E-01	AA431833.1	EST_HUMAN	zw79703.1 Soares testis_NHT Homo sapiens cDNA clone IMAGE:782429 5' similar to TR:G1066835
6593	19653	32925	0.68	3.5E-01	U37150.1	NT	G1066835 F10F2.1;
6821	19875	33184	0.82	3.5E-01	O24357	SWISSPROT	Bos taurus peptide methionine sulfoxide reductase (msrA) mRNA, complete cds
7257	19892		3.65	3.5E-01	X98505.1	NT	GLUCOSE-6-PHOSPHATE 1-DEHYDROGENASE, CHLOROPLAST PRECURSOR (G6PD)
7789	20742	34115	0.81	3.5E-01	P47281	SWISSPROT	S. scrofa mRNA for CD31 protein (PECAM-1)
7789	20742	34116	0.81	3.5E-01	P47281	SWISSPROT	HISTIDYL-TRNA SYNTHETASE (HISTIDINE-TRNA LIGASE) (HISRS)
8023	20960	34356	0.7	3.5E-01	X03091.1	NT	HISTIDYL-TRNA SYNTHETASE (HISTIDINE-TRNA LIGASE) (HISRS)
8406	21375		2.17	3.5E-01	11448042	NT	E. coli L-arabinose transport operon with genes araF, araG and araH
8409	21378	34784	0.67	3.5E-01	BF358871.1	EST_HUMAN	Homo sapiens tumor protein p53-binding protein, 2 (TP53BP2), mRNA
8810	21777		0.58	3.5E-01	AF051561.1	NT	RC4-ET0024-260600-014-d07 ET0024 Homo sapiens cDNA
9279	22245	35674	1.14	3.5E-01	4507610	NT	Rattus norvegicus Na-K-Cl cotransporter (Nkcc1) mRNA, complete cds
10083	23019	36494	6.94	3.5E-01	Q07294	SWISSPROT	Homo sapiens tyrosine kinase non-receptor 1 (TNK1), mRNA
10246	23171	36680	5.51	3.5E-01	Z26825.1	NT	VOLTAGE-DEPENDENT N-TYPE CALCIUM CHANNEL ALPHA-1B SUBUNIT (CALCIUM CHANNEL, L TYPE, ALPHA-1 POLYPEPTIDE ISOFORM 5) (BRAIN CALCIUM CHANNEL III) (BIII)
10326	23250	36729	1.01	3.5E-01	BE174794.1	EST_HUMAN	X. laevis gene for albumin including HP1 enhancer
11086	24047	37569	2.48	3.5E-01	X61084.1	NT	QV2-HT0577-090400-128-c07 HT0577 Homo sapiens cDNA
11362	24311	37838	1.68	3.5E-01	AJ243178.1	NT	C. griseus rhodopsin gene for opsin protein
11362	24311	37839	1.68	3.5E-01	AJ243178.1	NT	Gallus gallus SPARC gene for osteonectin, promoter and exon 1
11919	24800	38391	1.67	3.5E-01	N77597.1	EST_HUMAN	Gallus gallus SPARC gene for osteonectin, promoter and exon 1
11940	24820		1.77	3.5E-01	M82885.1	NT	y280112.1 Soares multiple sclerosis_2NblHMSF Homo sapiens cDNA clone IMAGE:290375 5'
11991	24868	38463	1.6	3.5E-01	U05145.1	NT	Drosophila melanogaster dual bar protein (BarH2) gene, exon 1
12269	25991		1.8	3.5E-01	AF287488.1	NT	Human glucokinase (GCK) gene, repeat polymorphism
12341	25139		1.31	3.5E-01	X64585.1	NT	Schistosoma mansoni strain NMRI chromatin assembly factor 1 small subunit-like protein (RBAP48) mRNA, complete cds
12501	25240		2.56	3.5E-01	AE001774.1	NT	B. taurus alpA1 gene for F(0)F(1) ATP synthase alpha-subunit
13085	25842	31430	3.37	3.5E-01	H80814.1	EST_HUMAN	Thermoboga maritima section 88 of 136 of the complete genome
13085	25842	31431	3.37	3.5E-01	H80814.1	EST_HUMAN	ys94f1.1.1 Soares retina N2b4HR Homo sapiens cDNA clone IMAGE:218597 5'
							ys94f1.1.1 Soares retina N2b4HR Homo sapiens cDNA clone IMAGE:218597 5'

Page 68 of 546
Table 4

Single Exon Probes Expressed in Bone Marrow

Probe SEQ ID NO:	Exon SEQ ID NO:	ORF SEQ ID NO:	Expression Signal	Most Similar (Top) Hit BLAST E Value	Top Hit Accession No.	Top Hit Database Source	Top Hit Descriptor
13106	25627		1.67	3.5E-01	4758297	NT	Homo sapiens v-erb-b2 avian erythroblastic leukemia viral oncogene homolog 2 (neurofiblastoma derived oncogene homolog) (ERBB2), mRNA
708	19770		1.97	3.4E-01	AL242958.1	NT	Homo sapiens partial N-myc (exon 3), HPV45 L2, HPV45 L1, HPV45 E6, HPV45 E7 and HPV45 E1 genes isolated from IC4 cervical carcinoma cell line
977	14028	26882	9.08	3.4E-01	Y09788.2	NT	Pseudomonas fluorescens colR, colS genes, orf222 and partial lnaA gene
1328	14364	27332	2.78	3.4E-01	Y00554.1	NT	Azotobacter vinelandii nifA gene for NifA protein (positive regulatory element)
2410	15417	28441	2.01	3.4E-01	D80809.1	NT	Synechocystis sp. PCC6803 complete genome, 11/27, 1311235-1430418
3014	16072	28992	0.86	3.4E-01	AL163210.2	NT	Homo sapiens chromosome 21 segment HS21C010
3014	16072	28993	0.86	3.4E-01	AL163210.2	NT	Homo sapiens chromosome 21 segment HS21C010
3175	16230	28146	6.62	3.4E-01	U69305.1	NT	Canis familiaris rod photoreceptor cGMP-gated channel alpha-subunit (CNIGC1) mRNA, complete cds
3354	16406	28328	0.94	3.4E-01	AF034882.1	NT	Homo sapiens pulmonary surfactant protein D, promoter region and exon 1
3542	16588	28512	3.42	3.4E-01	AF106835.1	NT	Methylovorus sp. strain SS1 putative GrpE (grpE), DnaK (dnaK), and putative DnaJ (dnaJ) genes, complete cds
3804	16844		1.78	3.4E-01	BF446010.1	EST_HUMAN	7n9a01.x1 NCI_CGAP_Ov18 Homo sapiens cDNA clone IMAGE:3572232 3' similar to TR:Q8UJ15
4082	17116		1.48	3.4E-01	AA594196.1	EST_HUMAN	Q8UJ15.DJ1808.1 ;
4674	17695	30582	1.72	3.4E-01	BE068912.1	EST_HUMAN	no11b10.s1 NCI_CGAP_Phe1 Homo sapiens cDNA clone IMAGE:1100347 3'
4989	18004		4.71	3.4E-01	AI240973.1	EST_HUMAN	MR4-BT0403-230200-202-c01 BT0403 Homo sapiens cDNA
5768	18860	32040	2.74	3.4E-01	AL101594.2	NT	q95c05.x1 NCI_CGAP_Kid3 Homo sapiens cDNA clone IMAGE:1867208 3' similar to contains Alu repetitive element;
5909	18906		5.14	3.4E-01	AA095313.1	EST_HUMAN	Arabidopsis thaliana DNA chromosome 4, config fragment No. 80
6122	19200		1.74	3.4E-01	L02971.1	NT	zn12d11.s1 Stratagene MNT neuron (#837233) Homo sapiens cDNA clone IMAGE:547221 3'
6146	19221	32451	0.8	3.4E-01	BE748912.1	EST_HUMAN	Echovirus 22 1AB, 1C, 1D, 2A, 2B, 2C, 3A, 3B, 3C, 3D proteins RNA, complete mature peptides and cde
6229	19303	32535	1.81	3.4E-01	AW204806.1	EST_HUMAN	0075718111 NIH_MGC_55 Homo sapiens cDNA clone IMAGE:3838828 3'
6362	19431	32674	1.78	3.4E-01	AL120544.1	EST_HUMAN	UI-HBH-acl-e-12-0-JL.s1 NCI_CGAP_Sub3 Homo sapiens cDNA clone IMAGE:2719582 3'
6907	19959		1.19	3.4E-01	N95225.1	EST_HUMAN	DKFZ761A249.t1 761 (synonym: hanny2) Homo sapiens cDNA clone DKFZp761A249.6
7135	20111	33424	1.14	3.4E-01	AI468082.1	EST_HUMAN	zb33e12.s1 Scores_fetal_Jung_NbHL19W Homo sapiens cDNA clone IMAGE:307942 3'
7261	19896	33293	0.61	3.4E-01	BF678702.1	EST_HUMAN	trn63g05.x1 NCI_CGAP_Brr25 Homo sapiens cDNA clone IMAGE:2162840 3' similar to gb:S37431
8298	21207	34986	0.54	3.4E-01	AE000463.1	NT	LAMININ RECEPTOR (HUMAN);
8579	21547	34986	0.55	3.4E-01	Y14930.1	NT	602085283F1 NIH_MGC_83 Homo sapiens cDNA clone IMAGE:4249365 5'
8832	21789		1.8	3.4E-01	AA337063.1	EST_HUMAN	Escherichia coli K-12 MG1655 section 383 of 400 of the complete genome
							Homo sapiens TORAV28 gene, allele A4, partial
							EST41765 Endometrial tumor Homo sapiens cDNA 5' end

Page 69 of 546
Table 4

Single Exon Probes Expressed in Bone Marrow

Probe SEQ ID NO:	Exon SEQ ID NO:	ORF SEQ ID NO:	Expression Signal	Most Similar (Top) Hit BLAST E Value	Top Hit Accession No.	Top Hit Database Source	Top Hit Descriptor
8909	21875	35301	0.64	3.4E-01	L04690.1	NT	Citellus griseus cholesterol 7-alpha-hydroxylase gene, complete cds
9204	22170	36600	1.89	3.4E-01	9633624	NT	Bovine enterovirus strain K2577, complete genome
9567	22529	35978	4.43	3.4E-01	P26013	SWISSPROT	INTEGRIN BETA-8 PRECURSOR
9567	22529	35979	4.43	3.4E-01	P26013	SWISSPROT	INTEGRIN BETA-8 PRECURSOR
9776	22717		0.49	3.4E-01	AB017510.1	NT	Ephydiala fluviatilis mRNA for PLC-gammaS, complete cds
9801	21124	34527	6.03	3.4E-01	U19482.1	NT	Saccharomyces cerevisiae Marfip (MAF1) gene, complete cds
9801	21124	34528	6.03	3.4E-01	U19482.1	NT	Saccharomyces cerevisiae Marfip (MAF1) gene, complete cds
9855	22791	36243	0.43	3.4E-01	AF193857.1	NT	Dicrostelium discoidium putative CMF receptor CMFR1 mRNA, complete cds
10054	22981	38449	1.09	3.4E-01	U68763.1	NT	Glycine max putative transcription factor SCOF-1 (scot-1) mRNA, complete cds
10249	23174	36984	2.14	3.4E-01	AJ225084.1	NT	Homo sapiens FAA gene, exon 16, 17 and 18
10843	23763		0.68	3.4E-01	AE004096.1	NT	Vibrio cholerae chromosome I, section 4 of 251 of the complete chromosome
11357	24307		3.51	3.4E-01	AE000891.1	NT	Methanobacterium thermoautotrophicum from bases 1018444 to 1028212 (section 87 of 148) of the complete genome
11393	24339	37669	4.96	3.4E-01	P09625	SWISSPROT	PROBABLE E4 PROTEIN
11427	24371	37909	1.67	3.4E-01	AF046881.1	NT	Rutillus ercaasi cytochrome b (cytb) gene, mitochondrial gene encoding mitochondrial protein, partial cds
11828	24711	38295	1.58	3.4E-01	AB035507.1	NT	Rattus norvegicus mRNA for s-glycerol/MUC18, complete cds
11856	24738	38323	3.3	3.4E-01	AL161515.2	NT	Arabidopsis thaliana DNA chromosome 4, contig fragment No. 27
12089	24970	38557	1.9	3.4E-01	AJ559886.1	EST_HUMAN	ta77g08.x1 NCL_CGAP_UH1 Homo sapiens cDNA clone IMAGE:2214874 3' similar to contains L1.b1 L1 repetitive element;
12149	25008		2.08	3.4E-01	U63604.1	NT	Citrus variegation virus putative replicase gene, partial cds
12261	25085		1.59	3.4E-01	Z21621.1	NT	S.cerevisiae RIB5 gene encoding Riboflavin synthase
12485	25231		13.04	3.4E-01	L26339.1	NT	Human autoantigen mRNA, complete cds
12512	25761		3.88	3.4E-01	BE218652.1	EST_HUMAN	h42408.x1 NCL_CGAP_Lu24 Homo sapiens cDNA clone IMAGE:3176127 3' similar to contains PTR5.13
12567	25965		2.19	3.4E-01	9639361	NT	PTR5 repetitive element;
12673	25345	31763	3.46	3.4E-01	AJ287131.1	NT	Beta vulgaris mitochondrion, complete genome
						NT	Mus musculus SIL, MAP_17, CYP_a, SCL & CYP_b genes
						NT	Homo sapiens HLA class III region containing tenascin X (tenascin-X) gene, partial cds; cytochrome P450 21-hydroxylase (CYP21B), complement component C4 (C4B) G11, helicase (SK12W), RD, complement factor B (Bf), and complement component C2 (C2) genes;>
12672	25538		1.94	3.4E-01	AF019413.1	NT	Rhizobium leguminosarum sym plasmid pRL5J1 nodX gene
15	13135	26033	10.37	3.3E-01	X07890.1	NT	Rhizobium leguminosarum sym plasmid pRL5J1 nodX gene
106	13135	26033	4.34	3.3E-01	X07890.1	NT	Rhizobium leguminosarum sym plasmid pRL5J1 nodX gene
448	13521	26454	1.3	3.3E-01	AL161545.2	NT	Arabidopsis thaliana DNA chromosome 4, contig fragment No. 45
634	13700	26621	2.28	3.3E-01	7662485	NT	Homo sapiens KIAA1100 protein (KIAA1100), mRNA

Page 70 of 546
Table 4

Single Exon Probes Expressed In Bone Marrow

Probe SEQ ID NO:	Exon SEQ ID NO:	ORF SEQ ID NO:	Expression Signal	Most Similar (Top) Hit BLAST E Value	Top Hit Accession No.	Top Hit Database Source	Top Hit Descriptor
1205	14244	27203	3.29	3.3E-01	Q12446	SWISSPROT	PROLINE-RICH PROTEIN LAS17
1310	14346	27312	2.44	3.3E-01	BF568880.1	EST_HUMAN	602184016T1 NIH_MGC_42 Homo sapiens cDNA clone IMAGE:4300261 3'
1610	14642	27618	1.55	3.3E-01	6753686	NT	Mus musculus disintegrin 5 (Dign5), mRNA
1752	14781		1.13	3.3E-01	AA332734.1	EST_HUMAN	EST367722 Embryo, 8 week Homo sapiens cDNA 5' end
2043	15062		1.01	3.3E-01	AF031148.1	NT	Methylococcus capsulatus strain Bath outer membrane protein MopB (mopB) gene, complete cds
2414	15421		4.45	3.3E-01	4507834	NT	Homo sapiens uridine monophosphate synthetase (crotate phosphoribosyl transferase and orididine-5'-decarboxylase) (UMPS) mRNA
2860	16018	28945	1.76	3.3E-01	AJ251805.1	NT	Bacteriophage phi-Yeo3-12 complete genome
3028	16086		0.8	3.3E-01	O02743	SWISSPROT	INTERLEUKIN-12 ALPHA CHAIN PRECURSOR (IL-12A) (CYTOTOXIC LYMPHOCYTE MATURATION FACTOR 35 KD SUBUNIT) (CLMF P38)
3068	16126	29038	0.91	3.3E-01	AJ007932.2	NT	Streptomyces argillaceus mitramycin biosynthetic genes
3506	16553	29479	1.27	3.3E-01	AB012922.1	NT	Homo sapiens MTA1-L1 gene, complete cds
3822	16862	29766	2.18	3.3E-01	O84645	SWISSPROT	EXODEOXYRIBONUCLEASE V BETA CHAIN
3832	16872	29773	0.8	3.3E-01	P22602	SWISSPROT	GENOME POLYPROTEIN [CONTAINS: N-TERMINAL PROTEIN (P1); HELPER COMPONENT PROTEINASE (HC-PRO); PROTEIN P3]
3989	17028	29839	1.54	3.3E-01	AL161498.2	NT	Arabidopsis thaliana DNA chromosome 4, contig fragment No. 10
4026	17084	29865	2.02	3.3E-01	AF200446.1	NT	Hypoxylon fragiforme chitin synthase gene, partial cds
4395	17423		1.51	3.3E-01	D31662.1	NT	Rattus norvegicus DNA for regucalcin, partial cds
4715	17735		1.41	3.3E-01	AJ539114.1	EST_HUMAN	tp78b12.x1 NCI_CGAP_UK3 Homo sapiens cDNA clone IMAGE:2205407 3' similar to gb:X57522 ANTIGEN PEPTIDE TRANSPORTER 1 (HUMAN);
4875	17892	30781	1.33	3.3E-01	D64003.1	NT	Synechocystis sp. PCC6803 complete genome, 22/27, 2755703-2868766
5397	18500	31377	2.48	3.3E-01	X89819.1	NT	R.norvegicus mRNA for 3'UTR of ubiquitin-like protein
5397	18500	31378	2.48	3.3E-01	X89819.1	NT	R.norvegicus mRNA for 3'UTR of ubiquitin-like protein
5684	18759	31928	0.55	3.3E-01	P39055	SWISSPROT	DYNAMIN
5684	18759	31929	0.55	3.3E-01	P39055	SWISSPROT	DYNAMIN
5884	18973	32166	0.61	3.3E-01	BF213873.1	EST_HUMAN	601848090F1 NIH_MGC_55 Homo sapiens cDNA clone IMAGE:4078823 5'
6056	19137	32347	1.75	3.3E-01	BE619650.1	EST_HUMAN	601472768T1 NIH_MGC_68 Homo sapiens cDNA clone IMAGE:3875753 3'
6056	19137	32348	1.75	3.3E-01	BE619650.1	EST_HUMAN	601472768T1 NIH_MGC_68 Homo sapiens cDNA clone IMAGE:3875753 3'
6154	19229	32459	0.82	3.3E-01	P05691	SWISSPROT	CIRCUMSPOROZOITE PROTEIN (CS)
6960	20185	33508	0.63	3.3E-01	AB034233.1	NT	Flexibacter litoralis gyrB gene for DNA gyrase B subunit, partial cds
6960	20185	33509	0.63	3.3E-01	AB034233.1	NT	Flexibacter litoralis gyrB gene for DNA gyrase B subunit, partial cds
7073	20086	33404	4.16	3.3E-01	AJ628131.1	EST_HUMAN	tp44p01.x1 NCI_CGAP_UK11 Homo sapiens cDNA clone IMAGE:2285809 3' similar to contains Alu repetitive element; contains element L1 repetitive element;

Table 4

Single Exon Probes Expressed in Bone Marrow

Probe SEQ ID NO:	Exon SEQ ID NO:	ORF SEQ ID NO:	Expression Signal	Most Similar (Top) Hit BLAST E Value	Top Hit Accession No.	Top Hit Database Source	Top Hit Descriptor
7073	20095	33405	4.16	3.3E-01	AI828131.1	EST_HUMAN	ly64h01.x1 NCL CGAP_Kid11 Homo sapiens cDNA clone IMAGE:2285809 3' similar to contains Alu repetitive element; contains element L1 repetitive element;
8082	20899	34395	1.81	3.3E-01	N85146.1	EST_HUMAN	J2498F Human fetal heart, Lambda ZAP Express Homo sapiens cDNA clone J2498 5' similar to TEGT
8908	21874	35300	22.63	3.3E-01	BF683954.1	EST_HUMAN	602140372F1 NIH_MGC_48 Homo sapiens cDNA clone IMAGE:4301800 5'
9080	22048	35489	0.63	3.3E-01	BF210322.1	EST_HUMAN	601873281F1 NIH_MGC_54 Homo sapiens cDNA clone IMAGE:4097180 5'
9489	22433	35871	0.85	3.3E-01	Q62926	SWISSPROT	MITOGEN-ACTIVATED PROTEIN KINASE KINASE 1 (MAPK/ERK KINASE 1) (MEK KINASE 1) (MEKK 1)
9733	22761	36215	0.98	3.3E-01	BE828461.1	EST_HUMAN	CM3-ET0041-180500-187-d10 ET0041 Homo sapiens cDNA
9733	22761	36216	0.98	3.3E-01	BE828461.1	EST_HUMAN	CM3-ET0041-180500-187-d10 ET0041 Homo sapiens cDNA
9869	22805	36258	2.8	3.3E-01	N69898.1	EST_HUMAN	2a67h01.s1 Soares fetal lung NBHL19W Homo sapiens cDNA clone IMAGE:287849 3'
9910	22731	36186	2.61	3.3E-01	BF376746.1	EST_HUMAN	RC4-TN0077-250600-011-g04 TN0077 Homo sapiens cDNA
10350	23274		2.12	3.3E-01	L41044.1	NT	Homo sapiens high-mobility group phosphoprotein (HMG1-LC) gene, exons 1-3, complete cds
11076	24038	37561	2.85	3.3E-01	X63953.1	NT	D.mauritiana Adh gene
11076	24038	37562	2.85	3.3E-01	X63953.1	NT	D.mauritiana Adh gene
11365	24313		2.16	3.3E-01	BF526469.1	EST_HUMAN	602070802F1 NCL CGAP_Bm04 Homo sapiens cDNA clone IMAGE:4213585 5'
11565	24505	38062	8.16	3.3E-01	BE218351.1	EST_HUMAN	nv51g02.x1 NCL CGAP_Lu24 Homo sapiens cDNA clone IMAGE:3176978 3'
11673	24639	38218	3.19	3.3E-01	P47953	SWISSPROT	GALECTIN-3 (GALACTOSE-SPECIFIC LECTIN 3) (MAC-2 ANTIGEN) (IGE-BINDING PROTEIN) (35 KD LECTIN) (CARBOHYDRATE BINDING PROTEIN 35) (CBP 35) (LAMININ-BINDING PROTEIN) (LECTIN L-29) (CBP30)
12027	24903		3.43	3.3E-01	AA808621.1	EST_HUMAN	db71g02.s1 NCL CGAP_GC81 Homo sapiens cDNA clone IMAGE:1336850 3'
12044	13135	26033	1.97	3.3E-01	X07980.1	NT	Rhizobium leguminosarum sym plasmid pRL5J1 nodX gene
12246	25075	38170	1.84	3.3E-01	6598319	NT	Homo sapiens aldehyde oxidase 1 (AOX1), mRNA
12668	25530		4.82	3.3E-01	AF000002.1	NT	Pyrococcus horikoshii OT3 genomic DNA, 287001-544000 nt, position (2/7)
13113	25632	31621	1.59	3.3E-01	BE312820.1	EST_HUMAN	601146730F1 NIH_MGC_19 Homo sapiens cDNA clone IMAGE:3162006 5'
457	13530		2.08	3.2E-01	AF018261.1	NT	Rattus norvegicus EH domain binding protein Epsin mRNA, complete cds
719	13781		0.76	3.2E-01	AL161661.2	NT	Arabidopsis thaliana DNA chromosome 4, contig fragment No. 61
1166	14207	27161	10.6	3.2E-01	AF047013.1	NT	Fusarium poae virus 1 RNA2 putative RNA dependent RNA polymerase gene, complete cds
1287	14322	27285	1.77	3.2E-01	Z50202.1	NT	P.vulgaris aro5-1 gene
1391	14425	27394	6.96	3.2E-01	Q48624	SWISSPROT	LACTOSE PERMEASE (LACTOSE-PROTON SYMPORT) (LACTOSE TRANSPORT PROTEIN)
1639	14671		0.9	3.2E-01	AF209730.1	NT	Arabidopsis thaliana cultivar Columbia RPP13 (RPP13) gene, complete cds
1789	14818	27803	1.5	3.2E-01	Z36041.1	NT	S. cerevisiae chromosome II reading frame ORF YBR172c
1799	14828	27815	5.47	3.2E-01	AW657194.1	EST_HUMAN	EST369284 IMAGE resequences, MAGD Homo sapiens cDNA
1799	14828	27816	6.47	3.2E-01	AW657194.1	EST_HUMAN	EST369284 IMAGE resequences, MAGD Homo sapiens cDNA

Table 4

Single Exon Probes Expressed in Bone Marrow

Probe SEQ ID NO:	Exon SEQ ID NO:	ORF SEQ ID NO:	Expression Signal	Most Similar (Top) Hit BLAST E Value	Top Hit Accession No.	Top Hit Database Source	Top Hit Descriptor
1830	14886	27882	1.03	3.2E-01	AL111655.1	NT	Botrytis cinerea strain T4 cDNA library under conditions of nitrogen deprivation
2168	15184	28205	2.33	3.2E-01	BF203817.1	EST_HUMAN	601888804F1 NIH_MGC_17 Homo sapiens cDNA clone IMAGE:4111612 5'
2548	15549		2.3	3.2E-01	7710079	NT	Mus musculus P04002011 homeobox (Pbx1), mRNA
2720	15714	28732	1.56	3.2E-01	AF080588.1	NT	Homo sapiens promyelocytic leukemia zinc finger protein (PLZF) gene, complete cds
3622	16665		0.79	3.2E-01	D10872.1	NT	Human h NAT allele 3-2 gene for arylamine N-acetyltransferase
4367	17394	30273	0.93	3.2E-01	4759105	NT	Homo sapiens symplekin (SYM) mRNA
4422	17449	30340	1.62	3.2E-01	M18818.1	NT	Rabbit beta-like globin gene cluster encoding the epsilon, gamma, delta (pseudogene) and beta globin polypeptides, complete cds
4526	17551	30439	1.3	3.2E-01	Q10268	SWISSPROT	HYPOTHETICAL 81.7 KD PROTEIN C13G7.04C IN CHROMOSOME 1 PRECURSOR
4767	17787		8.32	3.2E-01	BF683617.1	EST_HUMAN	602081872F1 NIH_MGC_81 Homo sapiens cDNA clone IMAGE:4246605 5'
4915	17932	30823	0.69	3.2E-01	Q57081	SWISSPROT	CYTADHERENCE HIGH MOLECULAR WEIGHT PROTEIN 3 (CYTADHERENCE ACCESSORY PROTEIN 3) (ACCESSORY ADHESIN PROTEIN 3) (P69)
5250	18258	31127	0.96	3.2E-01	AY008847.1	NT	Homo sapiens interleukin 12 p40 subunit (IL12B) gene, IL12B-1 allele, complete cds
5280	18286		4.18	3.2E-01	A888472.1	EST_HUMAN	ws2506.x1 NC1_CGAP_G08 Homo sapiens cDNA clone IMAGE:2488185 3' similar to contains Alu repetitive element; contains element PTR7 repetitive element;
5344	18449	31320	2.71	3.2E-01	BE173984.1	EST_HUMAN	CM0-H10569-060300-269-F10 HT0569 Homo sapiens cDNA
6068	19149	32361	1.36	3.2E-01	L27221.1	NT	Giardia intestinalis pyruvate:flavodoxin oxidoreductase and flanking genes
6436	19502	32763	0.68	3.2E-01	AF016494.1	NT	Fugu rubripes gamma-aminobutyric acid receptor beta subunit gene, partial cds; 55kd erythrocyte membrane protein (P65), synaptic vesicle-associated integral membrane protein (VAMP-1), procollagen C-protease enhancer protein (PCOLCE) genes, complete c>
6749	19803	33084	0.91	3.2E-01	AV718037.1	EST_HUMAN	AV718037 FHTA Homo sapiens cDNA clone FHTAABH01 5'
6897	19949		1.03	3.2E-01	AB002359.1	NT	Human mRNA for KIAA0381 gene, KIAA0381 protein
8189	21168	34668	0.44	3.2E-01	AJ277691.1	NT	Homo sapiens partial LMO1 gene for LIM domain only 1 protein, exon 1
8512	21480	34894	1.43	3.2E-01	M60268.1	NT	Rat ISO-ethyl natriuretic factor gene, complete cds
8609	21577	34993	0.45	3.2E-01	AJ231001.1	NT	Rattus norvegicus repeat; map NOS-D12Wax1
8710	21678	35103	16.12	3.2E-01	X02508.1	NT	H.sapiens gene fragment for acetylcholine receptor (AChR) alpha subunit exons 8, 9 and 3' flanking region
8713	21681	35108	17.12	3.2E-01	BF311635.1	EST_HUMAN	601897107F1 NIH_MGC_19 Homo sapiens cDNA clone IMAGE:4126633 5'
8805	21772		1.54	3.2E-01	AL161574.2	NT	Arabidopsis thaliana DNA chromosome 4, contig fragment No. 70
8845	21812	35231	1.13	3.2E-01	BF246771.1	EST_HUMAN	601855580F1 NIH_MGC_57 Homo sapiens cDNA clone IMAGE:4075627 5'
8845	21812	35232	1.13	3.2E-01	BF246771.1	EST_HUMAN	601855580F1 NIH_MGC_57 Homo sapiens cDNA clone IMAGE:4075627 5'
8919	21885	35311	2.13	3.2E-01	AE02015.1	NT	Deinococcus radiodurans R1 section 152 of 229 of the complete chromosome 1
9019	21895	35405	0.63	3.2E-01	U51026.1	NT	Oryctolagus cuniculus Ig H-chain pseudogene, V-region (VH6-e2) gene, partial cds
9019	21895	35406	0.63	3.2E-01	U51026.1	NT	Oryctolagus cuniculus Ig H-chain pseudogene, V-region (VH6-e2) gene, partial cds

Page 73 of 546
Table 4

Single Exon Probes Expressed in Bone Marrow

Probe SEQ ID NO:	Exon SEQ ID NO:	ORF SEQ ID NO:	Expression Signal	Most Similar (Top) Hit BLAST E Value	Top Hit Accession No.	Top Hit Database Source	Top Hit Descriptor
9422	22387	35826	0.49	3.2E-01	AL163204.2	NT	Homo sapiens chromosome 21 segment HS21C004
9432	22398		2.28	3.2E-01	M88511.1	NT	Human monocyte antigen CD14 (CD14) mRNA, complete cds
9605	22468	35811	0.45	3.2E-01	AF041828.1	NT	Homo sapiens 6-phosphofructo-2-kinase/fructose-2,6-bisphosphatase (PF2K) gene, exons 12 and 13
9605	22468	35812	0.45	3.2E-01	AF041828.1	NT	Homo sapiens 6-phosphofructo-2-kinase/fructose-2,6-bisphosphatase (PF2K) gene, exons 12 and 13
10553	23277	36752	3.37	3.2E-01	U44814.1	NT	Borrelia burgdorferi plasmid cp32-2, erpC and erpD genes, complete cds; and unknown genes
10559	23481	36976	0.5	3.2E-01	BE326230.1	EST_HUMAN	h98905.x1 NCI_CGAP_L124 Homo sapiens cDNA clone IMAGE:3181569 3'
10674	23598		3.71	3.2E-01	AB011389.1	NT	Homo sapiens gene for AF-6, complete cds
11028	23982	37519	3.03	3.2E-01	T06813.1	EST_HUMAN	EST04702 Fetal brain, Stratiene (cat#930206) Homo sapiens cDNA clone HFB221
12288	25890		3.67	3.2E-01	L07288.1	NT	Drosophila melanogaster laminin A (Lam-A) mRNA, complete cds
12674	25943		1.39	3.2E-01	BE88848.1	EST_HUMAN	601507820F1 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:3909532 5'
12804	25431		4.84	3.2E-01	O83217	SWISSPROT	ELONGATION FACTOR TU (EF-TU)
12880	25719		1.48	3.2E-01	AF157625.1	NT	Bos taurus Inositol 1,4,5-trisphosphate receptor type I mRNA, complete cds
12936	25516		1.57	3.2E-01	L39874.1	NT	Homo sapiens deoxydicholate deaminase gene, complete cds
13001	25925	31906	1.33	3.2E-01	BE388776.1	EST_HUMAN	601275480F1 NIH_MGC_20 Homo sapiens cDNA clone IMAGE:3616746 5'
13108	25628	31646	1.38	3.2E-01	AJ745111.1	EST_HUMAN	h21c06.x1 NCI_CGAP_Ov23 Homo sapiens cDNA clone IMAGE:2218954 3' similar to contains MER10.1
2679	15675	26396	3.02	3.1E-01	R18051.1	EST_HUMAN	MER10 repetitive element;
2708	15828	26717	4.84	3.1E-01	7881971	NT	ye90h08.r1 Soares fetal liver spleen 1NFLS Homo sapiens cDNA clone IMAGE:125051 5' similar to gb-M64241 QM PROTEIN (HUMAN);
2708	15828	26718	4.84	3.1E-01	7881971	NT	Homo sapiens KIAA0174 gene product (KIAA0174), mRNA
2689	15929		1.83	3.1E-01	AW620398.1	EST_HUMAN	Homo sapiens KIAA0174 gene product (KIAA0174), mRNA
3188	16243		4.03	3.1E-01	AB029069.1	NT	h48h08.x1 Soares_NFL_T_GBC_S1 Homo sapiens cDNA clone IMAGE:2875391 3'
3927	16967	28880	0.93	3.1E-01	AJ251588.1	NT	Mus musculus gene for Ser/Thr kinase KKIAMRE, exon 6
4999	18014	30901	1.25	3.1E-01	AE003984.1	NT	Daucus carota mRNA for transcription factor E2F (E2F gene)
5239	18247	31119	0.79	3.1E-01	AF130370.1	NT	Xylella fastidiosa, section 130 of 228 of the complete genome
5297	18320	31162	2.59	3.1E-01	AL163203.2	NT	Rattus norvegicus MEN1 tumor suppressor (Men1) mRNA, alternative splice product, complete cds
5555	18652	31598	9.78	3.1E-01	AF176114.1	NT	Homo sapiens chromosome 21 segment HS21C003
5681	18776	31948	0.65	3.1E-01	P44132	SWISSPROT	Homo sapiens hepatocyte nuclear factor-3 alpha (HNF3A) gene, exon 1
5682	18777	31949	0.87	3.1E-01	Z74883.1	NT	HYPOTHETICAL PROTEIN H1236
5693	18788		1.01	3.1E-01	Y13278.1	NT	S. cerevisiae chromosome XV reading frame ORF YOL141W
5866	18955	32142	2.3	3.1E-01	AF184122.1	NT	Mus musculus mRNA for polycystin
							Homo sapiens filamin 2 (FLN2) gene, exons 10 through 22

Page 74 of 546
Table 4

Single Exon Probes Expressed in Bone Marrow

Probe SEQ ID NO:	Exon SEQ ID NO:	ORF SEQ ID NO:	Expression Signal	Most Similar (Top) Hit BLAST E Value	Top Hit Accession No.	Top Hit Database Source	Top Hit Descriptor
6416	25660	32732	0.61	3.1E-01	R64322.1	EST_HUMAN	y441f04.1 Soares fetal liver spleen 1NRLS Homo sapiens cDNA clone IMAGE:198367 5'
6610	19668	32844	2.69	3.1E-01	AW963549.1	EST_HUMAN	RC3-HN0007-310300-011-504 HN0001 Homo sapiens cDNA
6683	19740	33016	0.85	3.1E-01	A1264458.1	EST_HUMAN	q39d01.x1 NCI CGAP_Co8 Homo sapiens cDNA clone IMAGE:1874689 3'
6840	19893	33188	0.88	3.1E-01	X71887.1	NT	H sapiens gene for immunoglobulin kappa light chain variable region A8 and A9
6830	20154		0.7	3.1E-01	AW377354.1	EST_HUMAN	MR2-CT02222-281099-005-H05 CT0222 Homo sapiens cDNA
7162	25634	31238	2.55	3.1E-01	BE737382.1	EST_HUMAN	601306121F1 NIH_MGC_39 Homo sapiens cDNA clone IMAGE:3640420 5'
7942	20884	34276	0.56	3.1E-01	4885390	NT	Homo sapiens hyaluronan synthase 2 (HAS2), mRNA
8038	20976	34371	0.49	3.1E-01	AF242431.1	NT	Mus musculus neuronal apoptosis inhibitory protein 6 (Naip6) gene, complete cds; and Naip3 gene, exons 2-6 and 11-18
8166	21104	34502	0.64	3.1E-01	AW850168.1	EST_HUMAN	IL3-CT0219-271089-022-E03 CT0219 Homo sapiens cDNA
8166	21104	34503	0.54	3.1E-01	AW850168.1	EST_HUMAN	IL3-CT0219-271089-022-E03 CT0219 Homo sapiens cDNA
8695	21961	35388	0.83	3.1E-01	R45318.1	EST_HUMAN	y946f01.a1 Soares Infant brain 1NIB Homo sapiens cDNA clone IMAGE:35639 3'
10262	23187	36671	0.62	3.1E-01	6878322	NT	Mus musculus phosphatidylinositol-4-phosphate 5-kinase, type 1 gamma (Pip5k1c), mRNA
10427	23349	36833	1	3.1E-01	BF686839.1	EST_HUMAN	602124743F1 NIH_MGC_56 Homo sapiens cDNA clone IMAGE:4281611 5'
10427	23349	36834	1	3.1E-01	BF686839.1	EST_HUMAN	602124743F1 NIH_MGC_56 Homo sapiens cDNA clone IMAGE:4281611 5'
10488	23410	36907	1.73	3.1E-01	A1244001.1	EST_HUMAN	q61e11.x1 NCI CGAP_K048 Homo sapiens cDNA clone IMAGE:1983980 3' similar to gb:S55700
10688	23688		0.56	3.1E-01	T55325.1	EST_HUMAN	HYDROXYMETHYLGLUTARYL-COA LYASE PRECURSOR (HUMAN);
11186	24142	37076	2.35	3.1E-01	BF216117.1	EST_HUMAN	y647h08.s1 Stragene fetal spleen (#937205) Homo sapiens cDNA clone IMAGE:74367 3' similar to similar to gb:M81036_ma2 HEMOGLOBIN GAMMA-A AND GAMMA-G CHAINS (HUMAN)
11864	24746	38328	2.3	3.1E-01	7682291	NT	601883592F1 NIH_MGC_57 Homo sapiens cDNA clone IMAGE:4095814 5'
12133	25002	38808	1.68	3.1E-01	AF048683.1	NT	Homo sapiens KIAA0764 gene product (KIAA0764), mRNA
12133	25002	38809	1.68	3.1E-01	AF048683.1	NT	Homo sapiens transcription factor forkhead-like 7 (FKHL7) gene, complete cds
12415	25188		1.57	3.1E-01	AF294308.1	NT	Homo sapiens transcription factor forkhead-like 7 (FKHL7) gene, complete cds
12451	25210		4.64	3.1E-01	AF304162.1	NT	Anolis opalinus isolate QS NADH dehydrogenase subunit 2 (ND2) gene, complete cds; mitochondrial gene for mitochondrial product
12682	25286		4.18	3.1E-01	AF196933.1	NT	Silvestrodon vitreum 40S ribosomal protein S11 mRNA, partial cds
12844	25523		3.39	3.1E-01	AF196778.1	NT	Homo sapiens membrane-bound aminopeptidase P (XNPEP2) gene, complete cds
73	15808	26112	1.78	3.0E-01	6755083	NT	Homo sapiens transcription factor KCHM enhancer 3, JM11 protein, JM4 protein, JM5 protein, T64 protein, JM10 protein, A4 differentiation-dependent protein, triple LIM domain protein 6, and synaptophysin genes, complete cds; and L-type calcium channel α^2
254	13361	26277	8.98	3.0E-01	AJ271735.1	NT	Mus musculus protein kinase C, epsilon (Pkc), mRNA
1227	14265	27222	2.16	3.0E-01	AW300400.1	EST_HUMAN	Homo sapiens Xq pseudautosomal region; segment 1/2
1508	14541	27512	5.96	3.0E-01	AJ106755.1	NT	xs63f08.x1 NCI CGAP_Kid11 Homo sapiens cDNA clone IMAGE:2774943 3'
						NT	Balaenoptera physalus gene encoding atrial natriuretic peptide

Page 75 of 546
Table 4

Single Exon Probes Expressed In Bone Marrow

Probe SEQ ID NO:	Exon SEQ ID NO:	ORF SEQ ID NO:	Expression Signal	Most Similar (Top) Hit BLAST E Value	Top Hit Accession No.	Top Hit Database Source	Top Hit Descriptor
2146	15182	28178	1.13	3.0E-01	AF237778.1	NT	Rattus norvegicus Ca ²⁺ /calmodulin-dependent protein kinase II, alpha subunit mRNA, 3' untranslated region
3226	16280		1.28	3.0E-01	AB030481.1	NT	Corynebacterium sp. ALY-1 atyPG gene for polygluturonate lyase, complete cds
3878	16818	29827	1.40	3.0E-01	AW617785.1	EST_HUMAN	PM1-ST0282-261189-001-g01 ST0282 Homo sapiens cDNA
3988	17037	28944	1.02	3.0E-01	AJ271736.1	NT	Homo sapiens Xq pseudautosomal region, segment 2/2
4541	17584	30451	2.17	3.0E-01	AJ006755.1	NT	Balanoptera physalus gene encoding atrial natriuretic peptide
5425	18528	31407	5.19	3.0E-01	BE741629.1	EST_HUMAN	601584960F1 NIH_MGC_9 Homo sapiens cDNA clone IMAGE:3948734 5'
5506	18606	31536	0.58	3.0E-01	AF224688.1	NT	Homo sapiens mannosidase, beta A, lysosomal (MANBA) gene, and ubiquitin-conjugating enzyme E2D 3 (UBE2D3) genes, complete cds
5510	18610	31541	0.78	3.0E-01	AF228247.1	NT	Canigalo orthopoxvirus hemagglutinin gene, complete cds
5582	18678	31641	3.81	3.0E-01	BE683575.1	EST_HUMAN	RC3-BT0333-180700-111-403 BT0333 Homo sapiens cDNA
5582	18678	31642	3.81	3.0E-01	BE683575.1	EST_HUMAN	RC3-BT0333-180700-111-403 BT0333 Homo sapiens cDNA
5619	18716	31874	4.05	3.0E-01	U01247.1	NT	Mus musculus 129/sv Clara cell 10 kd protein (mCC10) gene, complete cds
7005	20131	33446	2.86	3.0E-01	D16313.1	NT	Mouse cytokerafin 15 gene, complete cds
7041	18373	31261	0.71	3.0E-01	U02369.1	NT	Strongylocentrotus purpuratus 34/67 kDa laminin-binding protein mRNA, partial cds
7111	20045	33347	0.98	3.0E-01	AF228247.1	NT	Canigalo orthopoxvirus hemagglutinin gene, complete cds
7195	20219	33551	0.59	3.0E-01	X63941.1	NT	S. Cerevisiae GAC1
7328	20299	33643	0.82	3.0E-01	AL163206.2	NT	Homo sapiens chromosome 21 segment HS21C008
7549	20512	33870	4.68	3.0E-01	10947007	NT	Mus musculus midbrain (Widh-pending), mRNA
7744	20698	34064	1.33	3.0E-01	AF071810.1	NT	Streptococcus pneumoniae strain DBL5 PspA (pspA) gene, partial cds
8259	21228	34638	1.28	3.0E-01	AE001755.1	NT	Thermotoga maritima section 67 of 138 of the complete genome
8716	21684		4.13	3.0E-01	8910161	NT	Mus musculus C-type (calcium dependent, carbohydrate recognition domain) lectin, superfamily member 9 (Clec9e), mRNA
8819	21788	35210	1.34	3.0E-01	BE568083.1	EST_HUMAN	601339079F-1 NIH_MGC_53 Homo sapiens cDNA clone IMAGE:3681594 5'
9181	22147	35574	0.74	3.0E-01	AF141678.1	NT	Streptomyces sulfonolactans isopenicillin N synthase (pcbC) gene, partial cds
9223	22189		0.71	3.0E-01	7001685	NT	Homo sapiens DKFZP686M0122 protein (DKFZP686M0122), mRNA
9573	22535	35998	0.95	3.0E-01	AF220507.1	NT	Anabaena PCC7120 cytosine-specific DNA methyltransferase (dmnB) gene, complete cds; putative enhancin phosphate ribosyltransferase gene, partial cds; and unknown gene
9832	22659	36320	0.51	3.0E-01	P76389	SWISSPROT	HYPOTHETICAL 59.5 KD PROTEIN IN WZA-ASMA INTERGENIC REGION
10327	23251	36730	0.73	3.0E-01	BF574612.1	EST_HUMAN	602193271F1 NIH_MGC_91-Homo sapiens cDNA clone IMAGE:4288336 5'
10501	23423	36822	0.47	3.0E-01	AF152598.3	NT	Actinobacillus actinomycetomycetans TadA (tadA), TadB (tadB), TadC (tadC), TadD (tadD), TadE (tadE), TadF (tadF), and TadG (tadG) genes, complete cds
10501	23423	36823	0.47	3.0E-01	AF152598.3	NT	Actinobacillus actinomycetomycetans TadA (tadA), TadB (tadB), TadC (tadC), TadD (tadD), TadE (tadE), TadF (tadF), and TadG (tadG) genes, complete cds

Table 4

Single Exon Probes Expressed In Bone Marrow

Probe SEQ ID NO:	Exon SEQ ID NO:	ORF SEQ ID NO:	Expression Signal	Most Similar (Top) Hit BLAST E Value	Top Hit Accession No.	Top Hit Database Source	Top Hit Descriptor
10760	23681	37177	0.85	3.0E-01	AW11811.1	EST_HUMAN	pxo3d10.x1 Soares_NFL_T_GBC_S1 Homo sapiens cDNA clone IMAGE:2606035 3'
10762	23683	37179	2.14	3.0E-01	AB030231.1	NT	Aspergillus oryzae bpa gene for ER chaperone BiP, complete cds
10782	23703	37201	0.82	3.0E-01	BF683841.1	EST_HUMAN	602140133F1 NIH_MGC_48 Homo sapiens cDNA clone IMAGE:4301087 5'
10782	23703	37202	0.82	3.0E-01	BF683841.1	EST_HUMAN	602140133F1 NIH_MGC_48 Homo sapiens cDNA clone IMAGE:4301087 5'
12061	24634	38528	2.5	3.0E-01	H51028.1	EST_HUMAN	yp84b10.l1 Soares fetal liver spleen 1NLS Homo sapiens cDNA clone IMAGE:194107 5'
12061	24634	38530	2.5	3.0E-01	H51028.1	EST_HUMAN	yp84b10.l1 Soares fetal liver spleen 1NLS Homo sapiens cDNA clone IMAGE:194107 5'
12686	25876		1.57	3.0E-01	AJ287631.1	NT	Rattus norvegicus mRNA for glyceraldehyde-3-phosphate dehydrogenase type 2 (gapdh-2 gene)
12879	25917		2.52	3.0E-01	6877786	NT	Mus musculus ribose 5-phosphate isomerase A (Rpia), mRNA
1747	14776		0.92	2.8E-01	AJ248896.1	NT	Mus musculus mas proto-oncogene and lgf2r gene for insulin-like growth factor type 2 and L41pe and Au76 pseudogenes
2039	15058	28077	1.19	2.8E-01	AE000736.1	NT	Aquifex acidicus section 88 of 109 of the complete genome
2260	15274	28288	1.2	2.8E-01	AF222718.1	NT	Chrysothymus synuroideus mitochondrion, complete genome
3285	16319	28240	1.92	2.8E-01	AW754239.1	EST_HUMAN	PM1-CT0328-171288-001-f12 CT0328 Homo sapiens cDNA
3285	16319	28241	1.92	2.8E-01	AW754239.1	EST_HUMAN	PM1-CT0328-171288-001-f12 CT0328 Homo sapiens cDNA
3912	16862	28863	1.03	2.8E-01	A610836.1	EST_HUMAN	tp21a11.x1 NCI_CGAP_Gas4 Homo sapiens cDNA clone IMAGE:2188412 3' similar to gb:D16050 NIL-2-A
3855	16995	28911	0.81	2.8E-01	A1789472.1	EST_HUMAN	ZINC FINGER PROTEIN (HUMAN) contains element L1 repetitive element;
4112	17146		0.7	2.8E-01	AW002802.1	EST_HUMAN	w14d10.x1 NCI_CGAP_Kid12 Homo sapiens cDNA clone IMAGE:2402803 3' similar to WP-C34F8.7
4511	17536	30420	1.24	2.8E-01	AA284468.1	EST_HUMAN	repetitive element;
4710	17731		1.5	2.8E-01	AL163207.2	NT	Homo sapiens chromosome 21 segment HS21C007
6156	18168		1.25	2.8E-01	A670889.1	EST_HUMAN	wa06103.x1 NCI_CGAP_Kid11 Homo sapiens cDNA clone IMAGE:2287308 3' similar to contains L1.12 L1
5275	18281	31145	3.3	2.8E-01	AJ131017.1	NT	repetitive element;
5277	18283	31146	1.12	2.8E-01	BE741980.1	EST_HUMAN	Mus musculus SCL gene locus
5330	18438		1.49	2.8E-01	R37485.1	EST_HUMAN	601594241F1 NIH_MGC_9 Homo sapiens cDNA clone IMAGE:3948178 5'
5469	20065	33372	0.75	2.8E-01	AF321001.1	NT	y77e12.s1 Soares infant brain 1NIB Homo sapiens cDNA clone IMAGE:28281 3'
5858	18948	32133	5.1	2.8E-01	X56098.1	NT	Susca domestica subsp. salina S-adenosylmethionine synthase 2 mRNA, complete cds
5858	18948	32134	5.1	2.8E-01	X56098.1	NT	B. subtilis levanase operon levD, levE, levF, levG and sacC (partial) genes for fructose phosphotransferase
5871	18960	32149	6.12	2.8E-01	6878682	NT	B. subtilis levanase operon levD, levE, levF, levG and sacC (partial) genes for fructose phosphotransferase
6174	18249	32482	1.35	2.8E-01	AA418145.1	EST_HUMAN	system polypeptides P16, P18, P28, P30 and levanase

Table 4

Single Exon Probes Expressed in Bone Marrow

Probe SEQ ID NO:	Exon SEQ ID NO:	ORF SEQ ID NO:	Expression Signal	Most Similar (Top) BLAST E Value	Top Hit Accession No.	Top Hit Database Source	Top Hit Descriptor
6412	19480	32727	1.05	2.9E-01	AI797128.1	EST_HUMAN	wc27c05.x1 NCI_CGAP_Lu24 Homo sapiens cDNA clone IMAGE:2342312 3' similar to contains L1.H1.L1 repetitive element;
6459	19524	32775	2.3	2.9E-01	U03420.1	NT	Bos taurus myosin I mRNA, complete cds
6599	19659	32831	0.6	2.9E-01	R69194.1	EST_HUMAN	y39c08.l1 Soares placenta Nb2HP Homo sapiens cDNA clone IMAGE:141615 5'
6599	19659	32832	0.6	2.9E-01	R69194.1	EST_HUMAN	y39c08.l1 Soares placenta Nb2HP Homo sapiens cDNA clone IMAGE:141615 5'
6683	19635		0.84	2.9E-01	Z50156.1	NT	D.discoidum gene for 34 kD actin binding protein
7043	20065	33372	0.58	2.9E-01	AF321001.1	NT	Suaeda maritima subsp. salina S-adenosylmethionine synthetase 2 mRNA, complete cds
7179	18410	31211	1.67	2.9E-01	AF142328.1	NT	Mus musculus Filth protein (Filth) gene, complete cds; and Lgln protein (Lgln) gene, partial cds
7303	20274	33610	2.99	2.9E-01	Q04399	SWISSPROT	PUTATIVE MULTICOPPER OXIDASE YDR506C
7367	20337	33687	1.68	2.9E-01	AF100958.1	NT	Mus musculus major histocompatibility locus class II region; Fas-binding protein Daxx (DAXX) gene, partial cds; Bng1 (BING1), tapasin (tapasin), RafGDS-like factor (RLF), KE2 (KE2), BING4 (BING4), beta1, 3-galactosyl transferase (beta1,3-galactosyl tr>
8252	21221	34630	1.67	2.9E-01	BE540422.1	EST_HUMAN	601065830F1 NIH_MGC_10 Homo sapiens cDNA clone IMAGE:3452287 5'
8262	21221	34631	1.67	2.9E-01	BE540422.1	EST_HUMAN	601065830F1 NIH_MGC_10 Homo sapiens cDNA clone IMAGE:3452287 5'
8489	21457	34974	0.49	2.9E-01	AJ237837.1	NT	Bos taurus partial stat5A gene, exons 5-19
8489	21457	34875	0.49	2.9E-01	AJ237837.1	NT	Bos taurus partial stat5A gene, exons 5-19
8502	21470		1.16	2.9E-01	BF217743.1	EST_HUMAN	Bos taurus partial stat5A gene, exons 5-19
8682	21650		0.45	2.9E-01	AF197456.1	NT	Buchnera aphidicola plasmid pLau isolate Mi 2-isopropylmalate synthase (leuA) gene, partial cds; 3-isopropylmalate dehydrogenase (leuB) gene, complete cds; and isopropylmalate dehydratase subunit (leuC) gene, partial cds
8942	21808	35333	0.77	2.9E-01	AU150910.1	EST_HUMAN	AU150910 NT2RP2 Homo sapiens cDNA clone NT2RP2003901 3'
9277	22243	35972	1.01	2.9E-01	AF225908.1	NT	Arabidopsis thaliana sulfonurea receptor-like protein mRNA, complete cds
9388	22353	35784	0.6	2.9E-01	M22452.1	NT	Baboon lymphocyte homing/adhesion receptor mRNA, complete cds
9601	22605	36053	0.76	2.9E-01	AJ248287.1	NT	Pyrococcus abyssi complete genome; segment 5/6
9601	22605	36054	0.76	2.9E-01	AJ248287.1	NT	Pyrococcus abyssi complete genome; segment 5/6
11240	24193	37711	1.7	2.9E-01	AF128943.1	NT	Trypanosoma cruzi stage-specific surface glycoprotein gp82 (gp82) mRNA, partial cds
11497	24440	37989	1.88	2.9E-01	V01394.1	NT	Torpedo californica mRNA encoding acetylcholine receptor gamma subunit
11497	24440	37990	1.88	2.9E-01	V01394.1	NT	Torpedo californica mRNA encoding acetylcholine receptor gamma subunit
11913	24794	38385	2.95	2.9E-01	AL138078.2	NT	Campylobacter jejuni NCTC11169 complete genome; segment 5/6
12109	24979	38579	2.1	2.9E-01	AW284742.1	EST_HUMAN	UI-H-BWO-arm-f-10-Q-U1.s1 NCI CGAP Sub8 Homo sapiens cDNA clone IMAGE:2729594 3'
12651	25334	31760	1.47	2.9E-01	AW005671.1	EST_HUMAN	wz88f05.x1 NCI_CGAP_Bm25 Homo sapiens cDNA clone IMAGE:2565921 3' similar to contains element
12728	25381	31747	4	2.9E-01	AF052453.1	NT	MER28 repetitive element;
13025	25573	31694	1.35	2.9E-01	V09037.1	NT	Homo sapiens TNF- α -inducible RNA binding protein (TIRP) gene, complete cds
						NT	Chlamydomonas reinhardtii mRNA for nitrite reductase structural locus

Page 78 of 546
Table 4

Single Exon Probes Expressed in Bone Marrow

Probe SEQ ID NO:	Exon SEQ ID NO:	ORF SEQ ID NO:	Expression Signal	Most Similar (Top) Hit BLAST E Value	Top Hit Accession No.	Top Hit Database Source	Top Hit Descriptor
13025	25573	31685	1.35	2.8E-01	Y08937.1	NT	Chlamydomonas reinhardtii mRNA for nitrite reductase structural locus
670	13640		2.2	2.8E-01	U67139.1	NT	Rattus norvegicus A-kinase anchoring protein AKAP150 mRNA, complete cds
576	13844		1.41	2.8E-01	L28145.1	NT	Prune dwarf virus movement protein, complete cds; coat protein, complete cds
1085	14129	27083	3.9	2.8E-01	AF168050.1	NT	Gaira guira oocyte maturation factor Mos (c-mos) gene, partial cds
1282	14317	27279	1.06	2.8E-01	BE313442.1	EST_HUMAN	601148733F1 NIH_MGC_19 Homo sapiens cDNA clone IMAGE:3163688 5'
1282	14317	27280	1.08	2.8E-01	BE313442.1	EST_HUMAN	601148733F1 NIH_MGC_19 Homo sapiens cDNA clone IMAGE:3163688 5'
1295	14330	27291	2.65	2.8E-01	D86560.1	NT	Human mRNA for serine/threonine protein kinase, complete cds
1740	14770	27756	2.22	2.8E-01	AW850020.1	EST_HUMAN	QV1-CT0364-120200-085-b05 CT0364 Homo sapiens cDNA
2028	15048	28082	2.08	2.8E-01	AL047620.1	EST_HUMAN	DKFZp58812321_r1 588 (synonym: huter1) Homo sapiens cDNA clone DKFZp58812321
2141	15158	28174	1.51	2.8E-01	AW671195.1	EST_HUMAN	h444b03.x1 Soares_NFL_T_GBC_S1 Homo sapiens cDNA clone IMAGE:2812333 3'
2479	15483	28507	2.18	2.8E-01	AE001494.1	NT	Escherichia coli K-12 MG1655 section 384 of 400 of the complete genome
2554	15556	28508	3.07	2.8E-01	AL161565.2	NT	Escherichia coli K-12 MG1655 section 384 of 400 of the complete genome
2674	15571	28689	1.35	2.8E-01	AB020876.1	NT	Arabidopsis thaliana DNA chromosome 4, contig fragment No. 65
2982	16040		1.48	2.8E-01	AF179480.1	NT	Arabidopsis thaliana mRNA for lipoyltransferase, complete cds
2983	16041	28983	2.57	2.8E-01	Z14037.1	NT	Toxoplasma gondii 90kDa heat-shock protein (HSP90) mRNA, partial cds
2983	16041	28984	2.57	2.8E-01	Z14037.1	NT	B. taurus microsatellite (ETH121)
3390	16439	29365	1.16	2.8E-01	AP000004.1	NT	B. taurus microsatellite (ETH121)
4021	17059	29860	1.72	2.8E-01	AE001180.1	NT	Pyrococcus horikoshii OT3 genomic DNA, 777001-994000 nt. position (47)
4150	17181		0.67	2.8E-01	AE004450.1	NT	Borrelia burgdorferi (section 66 of 70) of the complete genome
4226	17255		2.41	2.8E-01	A090868.1	EST_HUMAN	Pseudomonas aeruginosa PAO1, section 11 of 529 of the complete genome
4488	17513	30401	2.62	2.8E-01	P13615	SWISSPROT	ov44g10.x1 Soares_testis_NHT Homo sapiens cDNA clone IMAGE:1640226 3' similar to contains Alu repetitive element; contains element MER22 repetitive element
4822	17839	30737	0.92	2.8E-01	D15050.1	NT	RNA POLYMERASE BETA SUBUNIT (LARGE STRUCTURAL PROTEIN) (L PROTEIN)
4822	17839	30738	0.92	2.8E-01	D15050.1	NT	Human mRNA for transcription factor AREB8, complete cds
4886	17883	30771	2.71	2.8E-01	AF030154.1	NT	Human mRNA for transcription factor AREB8, complete cds
4887	17914	30804	1.37	2.8E-01	BF528188.1	EST_HUMAN	Bovine adenovirus 3 complete genome
4920	17937	30829	1.69	2.8E-01	A1272888.1	EST_HUMAN	602042601F1 NCI_CGAP_Bim67 Homo sapiens cDNA clone IMAGE:4180123 5'
5384	25637	31362	23.73	2.8E-01	AA349997.1	EST_HUMAN	q158c11.x1 Soares_NHIMPu_S1 Homo sapiens cDNA clone IMAGE:1876828 3' similar to contains Alu repetitive element; contains element LTR5 repetitive element
5687	18782	31954	2.52	2.8E-01	AB016625.1	NT	EST57072 Infant brain Homo sapiens cDNA 5' end
5915	19001		1.04	2.8E-01	AW992583.1	EST_HUMAN	Human sapiens OCTN2 gene, complete cds
6028	19111	32313	0.57	2.8E-01	AA765296.1	EST_HUMAN	GM1-BN0024-150200-118-g12 BN0024 Homo sapiens cDNA cso1d06.s1 NCI_CGAP_GCB1 Homo sapiens cDNA clone IMAGE:1303691 3' similar to gb:M94539 FK506-BINDING PROTEIN (HUMAN);

Page 79 of 546
Table 4
Single Exon Probes Expressed in Bone Marrow

Probe SEQ ID NO:	Exon SEQ ID NO:	ORF SEQ ID NO:	Expression Signal	Most Similar (Top) Hit BLAST E Value	Top Hit Accession No.	Top Hit Database Source	Top Hit Descriptor
6048	19129		0.49	2.8E-01	AA0404576.1	EST_HUMAN	z441901.r1 Soares ovary tumor NibHOT Homo sapiens cDNA clone IMAGE:724921 6' similar to contains Alu repetitive element
6300	25994		0.78	2.8E-01	M36668.1	NT	Bovine 680 bp repeated unit of 1.723 satellite DNA
6343	19412	32653	2.02	2.8E-01	AF003124.1	NT	Mesembryanthemum crystallinum fructose-bisphosphate aldolase mRNA, complete cds
6343	19412	32654	2.02	2.8E-01	AF003124.1	NT	Mesembryanthemum crystallinum fructose-bisphosphate aldolase mRNA, complete cds
6895	19947	93244	8.31	2.8E-01	BF611215.1	EST_HUMAN	U1-H-B14-ec-f-04-0-U1.s1 NCI_CGAP_Sub8 Homo sapiens cDNA clone IMAGE:3095182 3'
7201	20225	33557	0.52	2.8E-01	U68300.1	NT	Orthomyxa heterodius cytochrome b (cytb) gene, mitochondrial gene encoding mitochondrial protein, complete cds
7579	20541	33900	0.51	2.8E-01	BE861455.1	EST_HUMAN	601460157F1 NIH_MGC_68 Homo sapiens cDNA clone IMAGE:3892142 5'
7675	20633		1.03	2.8E-01	U09633.1	NT	Marsilea quadrifolia ribulose-1,6-bisphosphate carboxylase/oxygenase large subunit (rbcL) gene, chloroplast gene encoding chloroplast protein, partial cds
7724	20680	34044	0.77	2.8E-01	X66680.1	NT	L. esculentum y62 mRNA for GTP-binding protein
8429	21398	34808	1.12	2.8E-01	A1946126.1	EST_HUMAN	qp48h01.x1 NCI_CGAP_Co8 Homo sapiens cDNA clone IMAGE:1928289 3' similar to gb:U08323_cds1
8429	21398	34809	1.12	2.8E-01	A1946126.1	EST_HUMAN	MITOCHONDRIAL 60S RIBOSOMAL PROTEIN L3 (HUMAN);
8551	21519	34938	2.25	2.8E-01	U51688.1	NT	qp48h01.x1 NCI_CGAP_Co8 Homo sapiens cDNA clone IMAGE:1928289 3' similar to gb:U08323_cds1
8859	21826	35249	0.49	2.8E-01	AA911629.1	EST_HUMAN	MITOCHONDRIAL 60S RIBOSOMAL PROTEIN L3 (HUMAN);
8859	21802		7.34	2.8E-01	BF347847.1	EST_HUMAN	Homo sapiens lanosterol 14-alpha demethylase cytochrome P450 (CYP51) gene, exon 5
8821	22970	36127	0.91	2.8E-01	U17251.1	NT	6022h05.s1 NCI_CGAP_Co12 Homo sapiens cDNA clone IMAGE:1419893 3' similar to gb:M87789 IG
10069	22998		0.85	2.8E-01	L13654.1	NT	GAMMA-1 CHAIN C REGION (HUMAN);
10248	23173	36682	0.93	2.8E-01	AF132728.1	NT	602022887F1 NCI_CGAP_Bm67 Homo sapiens cDNA clone IMAGE:4158525 5'
10248	23173	36683	0.93	2.8E-01	AF132728.1	NT	Neurospora crassa negative regulator sulfur controller-2 (econ-2) gene, complete cds
10310	23234	36716	0.9	2.8E-01	AF284393.1	NT	Lycopodium obscurum peroxidase (TPX1) mRNA, complete cds
10420	23342	36828	4.98	2.8E-01	7709163	NT	Escherichia coli translocated intimin receptor Tir (tir) gene, complete cds
10676	23598		0.94	2.8E-01	8828164	NT	Escherichia coli translocated intimin receptor Tir (tir) gene, complete cds
10718	23640	37133	0.44	2.8E-01	BE959727.2	EST_HUMAN	Rattus norvegicus glycerol-3-phosphate dehydrogenase gene, promoters A and B and exons 1a and 1b; nuclear gene for mitochondrial product
11095	24055	37578	1.9	2.8E-01	BF241082.1	EST_HUMAN	Homo sapiens hypodermal protein (LOC51319), mRNA
11095	24055	37579	1.9	2.8E-01	BF241082.1	EST_HUMAN	Fujinami sarcoma virus, complete genome
11125	24085	37612	2.96	2.8E-01	BF686970.1	EST_HUMAN	601654822R1 NIH_MGC_57 Homo sapiens cDNA clone IMAGE:38339765 3'
11608	24546		4.02	2.8E-01	BF074023.1	EST_HUMAN	601880784F1 NIH_MGC_55 Homo sapiens cDNA clone IMAGE:4109360 6'
12312	25120	31843	1.39	2.8E-01	AF268477.1	NT	601880784F1 NIH_MGC_55 Homo sapiens cDNA clone IMAGE:4109350 5'
							601852148F1 NIH_MGC_56 Homo sapiens cDNA clone IMAGE:4076026 5'
							602137418F1 NIH_MGC_83 Homo sapiens cDNA clone IMAGE:4273853 5'
							Ovis aries tissue inhibitor of metalloproteinase 1 (TIMP1) gene, partial cds

Table 4

Single Exon Probes Expressed in Bone Marrow

Probe SEQ ID NO:	Exon SEQ ID NO:	ORF SEQ ID NO:	Expression Signal	Most Similar (Top) Hit BLAST E Value	Top Hit Accession No.	Top Hit Database Source	Top Hit Descriptor
12684	26361		23.64	2.8E-01	D83328.1	NT	Mus musculus DNA for prostaglandin D2 synthase, complete cds
12788	25420	31736	3.09	2.8E-01	BE178089.1	EST_HUMAN	PM4-HT0606-030400-001-407 HT0606 Homo sapiens cDNA
12817	25439	31742	1.37	2.8E-01	BE000116.1	EST_HUMAN	601673020F1 NIH_MGC_21 Homo sapiens cDNA clone IMAGE:39555996 5'
12869	25879		3.15	2.8E-01	11433829	NT	Homo sapiens CDC42-binding protein kinase beta (DMPK-like), mRNA
13087	25873		1.76	2.8E-01	AW025400.1	EST_HUMAN	w08905.x1 NCL_CGAP_Ki68 Homo sapiens cDNA clone IMAGE:2527828 3'
477	13549	28477	2.95	2.7E-01	Y17324.1	NT	Rattus norvegicus CDK104 mRNA
615	13680	26588	9.43	2.7E-01	AA450061.1	EST_HUMAN	z338b10.s1 Soares_fetus_Nib2-HF8_9w Homo sapiens cDNA clone IMAGE:788827 3' similar to
1269	14301	27262	2.18	2.7E-01	AB004906.1	NT	corinatus ALU repetitive element
1626	14668		2.21	2.7E-01	X79815.1	NT	Ipomoea purpurea transposable element Tip100 gene for transposase, complete cds
1742	14772	27757	2.88	2.7E-01	W58067.1	EST_HUMAN	G.lentibilia SR2 gene
1781	14820	27805	2.49	2.7E-01	P03341	SWISSPROT	z322h10.r1 Soares_fetus_Nib2-HF8 Homo sapiens cDNA clone IMAGE:341443 5'
2145	15885		2.29	2.7E-01	AF047575.1	NT	GAG POLYPROTEIN [CONTAINS: INNER COAT PROTEIN P12; CORE PROTEIN P16; CORE SHELL PROTEIN P30; NUCLEOPROTEIN P10]
2375	15383	28405	9.38	2.7E-01	Y13888.1	NT	Rattus norvegicus vesicular monamine transporter type 2, promoter region and exon 1
2461	15465	28488	3.97	2.7E-01	A1310658.1	EST_HUMAN	Feline immunodeficiency virus env gene, isolate ITT0088PIU (M89), partial
2809	15688	28891	1	2.7E-01	AF251276.1	NT	tax43c11.x2 NCL_CGAP_Lu25 Homo sapiens cDNA clone IMAGE:2046836 3' similar to contains element L1
2867	16055		0.91	2.7E-01	BF088284.1	EST_HUMAN	repetitive element:
4038	17076	28876	1.74	2.7E-01	A1928015.1	EST_HUMAN	Mus musculus serine protease inhibitor 14 (Spi14) mRNA, complete cds
4053	17080	28885	0.76	2.7E-01	AF216214.1	NT	CM1-HT0875-060900-385-e05 HT0875 Homo sapiens cDNA
4053	17090	28888	0.78	2.7E-01	AF216214.1	NT	w082611.x1 NCL_CGAP_Ki68 Homo sapiens cDNA clone IMAGE:2462828 3'
4081	17097	28891	2.33	2.7E-01	L77589.1	NT	Drosophila buzzatii alpha-esterase 6 (aE6) gene, partial cds
4928	17945	30837	0.76	2.7E-01	L27516.1	NT	Drosophila buzzatii alpha-esterase 6 (aE6) gene, partial cds
5079	18089		4.11	2.7E-01	AW856131.1	EST_HUMAN	Homo sapiens DiGeorge syndrome critical region, telomeric end
5339	18444	31197	2.13	2.7E-01	P17277	SWISSPROT	Triticum aestivum (Wcs66) gene, complete cds
5587	18564		1.11	2.7E-01	AB033171.1	NT	RC1-CT0286-230200-016-e03 CT0286 Homo sapiens cDNA
6476	18541	32787	0.51	2.7E-01	Q00818	SWISSPROT	HOMEBOX PROTEIN HOXA4 (CHOX-1.4)
6476	18541	32788	0.51	2.7E-01	Q00818	SWISSPROT	Asteropora mylaphthalma mitochondrial cytb gene for cytochrome b, partial cds
6765	18819	33100	1.06	2.7E-01	AE001094.1	NT	LATENT TRANSFORMING GROWTH FACTOR BETA BINDING PROTEIN 1 PRECURSOR
							(TRANSFORMING GROWTH FACTOR BETA-1 BINDING PROTEIN 1) (TGF-BETA1-BP-1)
							(TRANSFORMING GROWTH FACTOR BETA-1 MASKING PROTEIN, LARGE SUBUNIT)
							LATENT TRANSFORMING GROWTH FACTOR BETA BINDING PROTEIN 1 PRECURSOR
							(TRANSFORMING GROWTH FACTOR BETA-1 BINDING PROTEIN 1) (TGF-BETA1-BP-1)
							(TRANSFORMING GROWTH FACTOR BETA-1 MASKING PROTEIN, LARGE SUBUNIT)

Table 4

Single Exon Probes Expressed in Bone Marrow

Probe SEQ ID NO:	Exon SEQ ID NO:	ORF SEQ ID NO:	Expression Signal	Most Similar (Top) Hit BLAST E Value	Top Hit Accession No.	Top Hit Database Source	Top Hit Descriptor
6765	18819	33101	1.06	2.7E-01	AE001094.1	NT	Archaeoglobus fulgidus section 13 of 172 of the complete genome
6846	20170	33494	2.1	2.7E-01	Q61654	SWISSPROT	FIBRILLIN 1 PRECURSOR
7253	19888		0.84	2.7E-01	A1540070.1	EST_HUMAN	td08108.x1 NC1_CGAP_OCL1 Homo sapiens cDNA clone IMAGE:2075103 3'
7580	20542	33901	0.8	2.7E-01	Q11078	SWISSPROT	HYPOTHETICAL 20.9 KD PROTEIN B0563.3 IN CHROMOSOME X
7811	20760	34135	0.86	2.7E-01	Q01168	SWISSPROT	NITROGEN REGULATORY PROTEIN NUT1
7811	20780	34136	0.86	2.7E-01	Q01168	SWISSPROT	NITROGEN REGULATORY PROTEIN NUT1
7952	20883	34285	1.9	2.7E-01	AF248054.1	NT	Bos taurus micromolar calcium activated neutral protease 1 (CAPN1) gene, exons 11-20, and partial cds
7952	20883	34288	1.9	2.7E-01	AF248054.1	NT	Bos taurus micromolar calcium activated neutral protease 1 (CAPN1) gene, exons 11-20, and partial cds
8012	20950	34343	0.86	2.7E-01	AA351121.1	EST_HUMAN	EST58740 Infant brain Homo sapiens cDNA 5' end similar to similar to myosin-binding protein H
8012	20950	34344	0.86	2.7E-01	AA351121.1	EST_HUMAN	EST58740 Infant brain Homo sapiens cDNA 5' end similar to similar to myosin-binding protein H
8084	21021	34421	0.83	2.7E-01	L01081.1	NT	Oryctolagus cuniculus UDP-glucuronosyltransferase (UGT2B13) mRNA, complete cds
8198	21188	34578	0.83	2.7E-01	AA013147.1	EST_HUMAN	ze33511.1 Scores retina N2b-4-IR Homo sapiens cDNA clone IMAGE:360967 3' similar to contains Alu repetitive element
8360	21329		0.52	2.7E-01	AF049820.1	NT	Canis lupus familiaris pituitary adenylate cyclase activating polypeptide type 1 receptor precursor mRNA, complete cds
8475	21444	34861	0.43	2.7E-01	AW868503.1	EST_HUMAN	MRI-SN0062-100500-002-409 SN0062 Homo sapiens cDNA
8527	21495	34909	0.52	2.7E-01	R39257.1	EST_HUMAN	yc91h08.s1 Scores Infant brain 1N1B Homo sapiens cDNA clone IMAGE:23511 3'
8634	21602	35025	0.74	2.7E-01	AL161562.2	NT	Arabidopsis thaliana DNA chromosome 4, contig fragment No. 52
9112	22078	35505	0.52	2.7E-01	Q14764	SWISSPROT	MAJOR VAULT PROTEIN (MVP) (LUNG RESISTANCE-RELATED PROTEIN)
9382	22347	35779	0.51	2.7E-01	X03216.1	NT	Staphylococcus aureus transposon Tn554
9688	22841	36098	11.09	2.7E-01	O63809	SWISSPROT	THREONYL-TRNA SYNTHETASE (THREONINE-TRNA LIGASE) (THRRS)
9688	22841	36098	11.09	2.7E-01	O63809	SWISSPROT	THREONYL-TRNA SYNTHETASE (THREONINE-TRNA LIGASE) (THRRS)
9881	22844		2.17	2.7E-01	P37928	SWISSPROT	FIMBRIN W PROTEIN
10160	23085	36592	0.84	2.7E-01	D89880.1	NT	Rattus norvegicus DNA for peroxisome assembly factor-2, exon 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17 and complete cds
10441	23363	36853	0.85	2.7E-01	AF091848.1	NT	Oryctolagus cuniculus calgranulin C mRNA, partial cds
10477	23399	36898	2.83	2.7E-01	AF087434.1	NT	Mus musculus transcription factor NF-ATc isoform a (NF-ATc) mRNA, complete cds
10811	23533	37029	1.06	2.7E-01	AF156539.1	NT	Homo sapiens xeroderma pigmentosum complementation group C (XPC) gene, Intron 9
10811	23533	37030	1.06	2.7E-01	AF156539.1	NT	Homo sapiens xeroderma pigmentosum complementation group C (XPC) gene, Intron 9
11163	24121	37848	4.87	2.7E-01	AV705043.1	EST_HUMAN	AV705043 ADB Homo sapiens cDNA clone ADBCOD05 5'
11163	24121	37849	4.87	2.7E-01	AV705043.1	EST_HUMAN	AV705043 ADB Homo sapiens cDNA clone ADBCOD05 5'

Table 4

Single Exon Probes Expressed in Bone Marrow

Probe SEQ ID NO:	Exon SEQ ID NO:	ORF SEQ ID NO:	Expression Signal	Most Similar (Top) Hit BLAST E Value	Top Hit Accession No.	Top Hit Database Source	Top Hit Descriptor
11171	24128	37658	2.31	2.7E-01	AJ133269.1	NT	Homo sapiens caveolin-1/-2 locus, Cav1g1, D7S522, genes CAV2 (exons 1, 2a, and 2b), CAV1 (exons 1 and 2)
12069	24939		2.08	2.7E-01	BE141035.1	EST_HUMAN	MRO-HT0067-201089-002-c10 HT0067 Homo sapiens cDNA
12090	24961		1.52	2.7E-01	O14181	SWISSPROT	PUTATIVE 60S RIBOSOMAL PROTEIN C4F8.05C
12509	28244	31802	1.39	2.7E-01	AL163281.2	NT	Homo sapiens chromosome 21 segment HS21C081
12757	25758		1.46	2.7E-01	AB008782.1	NT	Arabidopsis thaliana mRNA for sulfate transporter, complete cds
12851	25460		1.63	2.7E-01	Q63627	SWISSPROT	CTD-BINDING SR-LIKE PROTEIN RA4
12848	25528		3.16	2.7E-01	AF217491.1	NT	Homo sapiens fragile 16D octo reductase (FOR) gene, exon 6
470	15941	28470	1.97	2.6E-01	P78411	SWISSPROT	IROQUOIS-CLASS HOMEODOMAIN PROTEIN IRX-2
481	13554		1.5	2.6E-01	D16459.1	NT	Bos taurus mRNA for mb-1, complete cds
1394	14428	27397	1.66	2.6E-01	BE885087.1	EST_HUMAN	601510838F1 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:3912345 5'
1431	14465	27441	1.36	2.6E-01	AB013280.1	NT	Glycine max pseudogene for Bd 30K
1912	14938	27831	6.4	2.6E-01	AL161472.2	NT	Arabidopsis thaliana DNA chromosome 4, contig fragment No. 2
1912	14938	27832	6.4	2.6E-01	AL161472.2	NT	Arabidopsis thaliana DNA chromosome 4, contig fragment No. 2
2104	15121		7.48	2.6E-01	AW733152.1	EST_HUMAN	bb04410.x1 NIH_MGC_14 Homo sapiens cDNA clone IMAGE:2888451 3' similar to gb:U36072 60S RIBOSOMAL PROTEIN L7A (HUMAN); gb:M14688_cds1 Mouse surfeit locus surfeit 3 protein gene (MOUSE);
2480	15484		0.99	2.6E-01	Y12998.1	NT	B. maritimus rbcL gene
2553	15555		8.24	2.6E-01	BE272440.1	EST_HUMAN	601128016F1 NIH_MGC_9 Homo sapiens cDNA clone IMAGE:2890043 5'
3108	19165		1.02	2.6E-01	AW974531.1	EST_HUMAN	EST386639 IMAGE resequences, MAGM Homo sapiens cDNA
3594	16639	28559	1.02	2.6E-01	M22342.1	NT	Bacteriophage T2 DNA (adenine-N6)methyltransferase (dam) gene, complete cds
3657	16700	28616	2.11	2.6E-01	AF229118.1	NT	Homo sapiens acetylcholinesterase collagen-like tail subunit (COLQ) gene, exons 1A, 2, 3, 4, and 5
3982	17022	28932	0.77	2.6E-01	AJ012174.2	NT	Chlamydomonas reinhardtii rnpB gene for RNase P RNA subunit
3982	17022	28933	0.77	2.6E-01	AJ012174.2	NT	Chlamydomonas reinhardtii rnpB gene for RNase P RNA subunit
4181	17212	30100	17.99	2.6E-01	BE080589.1	EST_HUMAN	QV1-BT0630-040400-132-e03 BT0630 Homo sapiens cDNA
4394	17412	30296	1.39	2.6E-01	AF175293.1	NT	Enterococcus faecium strain N97-330 vanD glycopeptide resistance gene cluster, complete cds; and unknown gene
4522	17547	30434	0.82	2.6E-01	AB021180.1	NT	Gallus gallus mRNA for skeletal myosin heavy chain, complete cds
4522	17547	30435	0.82	2.6E-01	AB021180.1	NT	Gallus gallus mRNA for skeletal myosin heavy chain, complete cds
4578	17601	30497	1.47	2.6E-01	AA457617.1	EST_HUMAN	ssod07.r1 Stratiotes fetal retina 637202 Homo sapiens cDNA clone IMAGE:339477 5'
4677	17698	30595	2.31	2.6E-01	U01103.1	NT	Arabidopsis thaliana PSI type III chlorophyll a/b-binding protein (Lhca3*) mRNA, complete cds
4745	17765	30659	1.46	2.6E-01	AF142703.1	NT	Ophrestia radicata maturase-like protein (matK) gene, complete cds; chloroplast gene for chloroplast product

Page 83 of 548
Table 4

Single Exon Probes Expressed in Bone Marrow

Probe SEQ ID NO:	Exon SEQ ID NO:	ORF SEQ ID NO:	Expression Signal	Most Similar (Top) Hit BLAST E Value	Top Hit Accession No.	Top Hit Database Source	Top Hit Descriptor
5007	18021	30809	4.38	2.6E-01	H04858.1	EST_HUMAN	yf51a05.1 Soares placenta NB2HP Homo sapiens cDNA clone IMAGE:152288 5'
5081	18091		0.78	2.6E-01	AA884625.1	EST_HUMAN	am33b11.s1 Soares_NFL_T_GBC_S1 Homo sapiens cDNA clone IMAGE:1468605 3'
5414	18517		1.06	2.6E-01	AB035972.1	NT	Paramoedium caudatum gene for PAP, complete cds
5524	18623	31538	0.69	2.6E-01	M80060.1	NT	Acetobacter xylinum cellulose synthase (bcsA) gene, partial cds, CMCax and CopAx genes, complete cds
5651	18747		0.71	2.6E-01	AI882398.1	EST_HUMAN	id18a03.x1 NCL_CGAP_Co18 Homo sapiens cDNA clone IMAGE:2075788 3' similar to contains element MER35 repetitive element;
5859	18958	32146	0.68	2.6E-01	AF207550.1	NT	Homo sapiens protein translocase, JM28 protein, UDP-galactose translocator, pim-2 protooncogene homolog pim-2h, and shal-type potassium channel genes, complete cds; JM12 protein and transcription factor IGHM enhancer 3 genes, partial cds; and unknown gp
6189	25993		2.61	2.6E-01	AE001811.1	NT	Thermotoga maritima section 123 of 138 of the complete genome
6325	18995	32837	1.81	2.6E-01	AI582557.1	EST_HUMAN	ta02a12.x1 NCL_CGAP_Pan1 Homo sapiens cDNA clone IMAGE:2227438 3' similar to SW:NDF1_RAT Q84288 NEUROGENIC DIFFERENTIATION FACTOR 1 ;contains element LTR1 repetitive element;
6325	18995	32838	1.81	2.6E-01	AI582557.1	EST_HUMAN	ta02a12.x1 NCL_CGAP_Pan1 Homo sapiens cDNA clone IMAGE:2227438 3' similar to SW:NDF1_RAT Q84288 NEUROGENIC DIFFERENTIATION FACTOR 1 ;contains element LTR1 repetitive element;
6562	19622	32887	1.01	2.6E-01	AL162757.2	NT	Neisseria meningitidis serogroup A strain Z2491 complete genome; segment 8/7
6825	19879	33169	0.8	2.6E-01	BE782052.1	EST_HUMAN	601581754F1 NIH_MGC_7 Homo sapiens cDNA clone IMAGE:3836156 5'
6825	19879	33170	0.8	2.6E-01	BE782052.1	EST_HUMAN	601581754F1 NIH_MGC_7 Homo sapiens cDNA clone IMAGE:3836156 5'
7238	20259	33593	0.75	2.6E-01	AB14380.1	EST_HUMAN	wd48c04.x1 Soares_NFL_T_GBC_S1 Homo sapiens cDNA clone IMAGE:2331368 3' similar to gp:M37721 PEPTIDYL-GLYCINE ALPHA-AMIDATING MONOOXYGENASE PRECURSOR (HUMAN);
7620	20580	33944	0.72	2.6E-01	BE148981.1	EST_HUMAN	GM0-HT0245-031189-086-04 HT0245 Homo sapiens cDNA
7683	25688		0.73	2.6E-01	AL139077.2	NT	Campylobacter jejuni NCTC11168 complete genome; segment 4/6
7700	20658		0.61	2.6E-01	AA196149.1	EST_HUMAN	zp82a01.11 Striatogene HeLa cell s3 937276 Homo sapiens cDNA clone IMAGE:627872 5'
8013	20851	34345	1.53	2.6E-01	R10385.1	EST_HUMAN	y937a03.s1 Soares fetal liver spleen 1NFLS Homo sapiens cDNA clone IMAGE:129004 3' similar to gb:U12517 U1 SMALL NUCLEAR RIBONUCLEOPROTEIN C (HUMAN);
8071	21008	34406	0.65	2.6E-01	Q09855	SWISSPROT	HYPOTHETICAL TRP-ASP REPEATS CONTAINING PROTEIN C28E6.01 IN CHROMOSOME I
8182	21152	34559	1.13	2.6E-01	R02411.1	EST_HUMAN	y982a07.1 Soares fetal liver spleen 1NFLS Homo sapiens cDNA clone IMAGE:124212 5'
8236	21205	34610	1.17	2.6E-01	BE144931.1	EST_HUMAN	MR0-HT0168-181189-003-412 HT0168 Homo sapiens cDNA
8490	21449	34866	0.82	2.6E-01	X82841.1	NT	D.melanogaster mRNA for alpha 1,2 mannosidase (Berlin)
8490	21449	34867	0.82	2.6E-01	X82841.1	NT	D.melanogaster mRNA for alpha 1,2 mannosidase (Berlin)
8677	21645	35069	2.87	2.6E-01	BF343588.1	EST_HUMAN	602014422F1 NCL_CGAP_Bm64 Homo sapiens cDNA clone IMAGE:4150386 5'
8753	21721	35143	1.92	2.6E-01	Q10189	SWISSPROT	HYPOTHETICAL 75.2 KD PROTEIN C11C1.02 IN CHROMOSOME II

Page 84 of 546
Table 4

Single Exon Probes Expressed in Bone Marrow

Probe SEQ ID NO.	Exon SEQ ID NO.	ORF SEQ ID NO.	Expression Signal	Most Similar (Top) Hit BLAST E Value	Top Hit Accession No.	Top Hit Database Source	Top Hit Descriptor
9042	22008	35428	4.09	2.6E-01	BE830339.1	EST_HUMAN	RC5-ET0082-310500-021-F10 ET0082 Homo sapiens cDNA
9042	22008	35429	4.09	2.6E-01	BE830339.1	EST_HUMAN	RC5-ET0082-310500-021-F10 ET0082 Homo sapiens cDNA
9822	22671	36128	0.99	2.6E-01	X17604.1	NT	S. occidentalis INV gene for Invertase (EC 3.2.1.26)
10098	23022		0.63	2.6E-01	AF057121.1	NT	Laritra canadensis cytochrome b (cytb) gene, mitochondrial gene encoding mitochondrial protein, complete cds
10227	23152	36841	1.08	2.6E-01	P87366	SWISSPROT	GREEN-SENSITIVE OPSIN (GREEN CONE PHOTORECEPTOR PIGMENT) (KHF-G)
10227	23152	36842	1.08	2.6E-01	P87366	SWISSPROT	GREEN-SENSITIVE OPSIN (GREEN CONE PHOTORECEPTOR PIGMENT) (KHF-G)
10550	23472		0.61	2.6E-01	Q28295	SWISSPROT	VON WILLEBRAND FACTOR PRECURSOR (VWF)
10875	23705		0.91	2.6E-01	Y10198.1	NT	Homo sapiens PHF1 gene
10878	23898		0.48	2.6E-01	Y16874.2	NT	Danio rerio mRNA for RPTP-alpha protein
11854	24736		32.27	2.6E-01	X51755.1	NT	Human lambda-immunoglobulin constant region complex (germline)
12268	25091		2.77	2.6E-01	10190655	NT	Mus musculus jerky (Jrk), mRNA
12464	25883		3.3	2.6E-01	BE883491.1	EST_HUMAN	801511052F1 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:3912612 5'
12528	25257	31808	4.42	2.6E-01	AF316898.1	NT	Homo sapiens NaK-ATPase gamma subunit (FXVD2) gene, complete cds, alternatively spliced
12846	25456		2.01	2.6E-01	D88425.1	NT	CaVia cabaia mRNA for serine/threonine kinase, complete cds
12874	25539		1.74	2.6E-01	AF141326.2	NT	Homo sapiens inositol polyphosphate 1-phosphatase (INPP1) gene, complete cds
13010	25562		3.07	2.6E-01	P47285	SWISSPROT	HYPOTHETICAL PROTEIN MG039
241	13340	26285	2.33	2.5E-01	4502298	NT	Homo sapiens ATP synthase, H+ transporting, mitochondrial F1 complex, delta subunit (ATP6D), nuclear gene encoding mitochondrial protein, mRNA
242	13340	26285	1.97	2.5E-01	4502298	NT	Homo sapiens ATP synthase, H+ transporting, mitochondrial F1 complex, delta subunit (ATP6D), nuclear gene encoding mitochondrial protein, mRNA
255	13362		4.63	2.5E-01	M26501.1	NT	Starfish (P. ochraceus) cytoplasmic actin gene, complete cds
833	13890	26844	1.97	2.5E-01	U09964.1	NT	Mus musculus ICR/Swiss glyceraldehyde 3-phosphate dehydrogenase (Gapd-S) gene, complete cds
1123	14167	27119	9.8	2.5E-01	T89837.1	EST_HUMAN	ye11g07.r1 Strategene lung (#837210) Homo sapiens cDNA clone IMAGE:117488 5'
1522	14554	27525	1.55	2.5E-01	AL115624.1	NT	Botrytis cinerea strain T4 cDNA library under conditions of nitrogen deprivation
1741	14771		5.06	2.5E-01	4885406	NT	Homo sapiens hyperpolarization activated cyclic nucleotide-gated potassium channel 4 (HCN4) mRNA
1900	15880	27920	0.94	2.5E-01	BE686804.1	EST_HUMAN	PM4-CT0400-310700-005-d08 CT0400 Homo sapiens cDNA
1900	15880	27921	0.94	2.5E-01	BE686804.1	EST_HUMAN	PM4-CT0400-310700-005-d08 CT0400 Homo sapiens cDNA
2417	15424		13.96	2.5E-01	AE000675.1	NT	Aquifex aeolicus section 7 of 109 of the complete genome
2503	15506		1.09	2.5E-01	AA251087.1	EST_HUMAN	2611a12.r1 NC1_CGAP_GCB1 Homo sapiens cDNA clone IMAGE:694862 5'
2644	15641	28665	0.94	2.5E-01	X96310.1	NT	B. taurus mRNA for D-aspartate oxidase
3423	16471		2.83	2.5E-01	AW973471.1	EST_HUMAN	EST385464 IMAGE: ressequences, MACM Homo sapiens cDNA

Page 85 of 546
Table 4
Single Exon Probes Expressed in Bone Marrow

Probe SEQ ID NO:	Exon SEQ ID NO:	ORF SEQ ID NO:	Expression Signal	Most Similar (Top) Hit BLAST E Value	Top Hit Accession No.	Top Hit Database Source	Top Hit Descriptor
3545	16591	29516	0.87	2.5E-01	AF233875.1	NT	Danio rerio peptide YY precursor gene, complete cds
3560	16806	29526	7.94	2.5E-01	AL161517.2	NT	Arabidopsis thaliana DNA chromosome 4, contig fragment No. 29
3801	16900	29803	1.15	2.5E-01	AI741483.1	EST_HUMAN	wg11c07.x1 Scores_NSF_F8_9W_OT_PA_P_S1 Homo sapiens cDNA clone IMAGE:2364780 3'
3861	16900	29804	1.15	2.5E-01	AI741483.1	EST_HUMAN	wg11c07.x1 Scores_NSF_F8_9W_OT_PA_P_S1 Homo sapiens cDNA clone IMAGE:2364780 3'
4096	17130		1.76	2.5E-01	P32323	SWISSPROT	A-AGGLUTININ ATTACHMENT SUBUNIT PRECURSOR
4348	17373		0.89	2.5E-01	Q03314	SWISSPROT	RHB PROTEIN
4780	17800		1.34	2.5E-01	Q27225	SWISSPROT	MOLT-INHIBITING HORMONE PRECURSOR (MIH)
4788	17808	30898	4.71	2.5E-01	AF007768.1	NT	Chlorostoma lumiferana diapause associated protein 2 (DAP2) mRNA, complete cds
4821	17838	30736	2.32	2.5E-01	AE004416.1	NT	Vibrio cholerae chromosome II, section 73 of 93 of the complete chromosome
4840	17857		3.89	2.5E-01	AJ230113.1	NT	Mus musculus annexin V gene, intron 4 segment containing 5' LTR and gag portion of MuERV-L (murine endogenous retrovirus) element
4870	17887	30775	0.84	2.5E-01	BE596785.1	EST_HUMAN	h62f11.x1 Scores_NFL_T_9BC_S1 Homo sapiens cDNA clone IMAGE:3041997 3' similar to WP:Y71F9A_294.D CE22658:
5086	18096	30972	0.84	2.5E-01	AW873598.1	EST_HUMAN	cc63a09.e1 NCI_CGAP_GCB1 Homo sapiens cDNA clone IMAGE:1316920 3' similar to contains AU repetitive element;
5243	18251		0.96	2.5E-01	AA768389.1	EST_HUMAN	T3 receptor-associated cofactor-1 [human, fetal liver, mRNA, 2830 nt]
5399	18502	31380	13.58	2.5E-01	S63390.1	NT	Homo sapiens KVLQ11 gene
6070	19151	32363	0.84	2.5E-01	AJ006345.1	NT	Homo sapiens chromosome 21 segment HS21C007
6071	19152		0.87	2.5E-01	AL163207.2	NT	PROTEIN KINASE YPS15
6525	19588	32847	0.49	2.5E-01	P22219	SWISSPROT	Homo sapiens partial siberin-1 gene
6760	19835	33118	0.86	2.5E-01	AJ251973.1	NT	Rattus norvegicus rebin 3 (RABIN3), mRNA
7245	19880	33277	0.8	2.5E-01	8394138	NT	Feline calicivirus CFV68 RNA helicase/cysteine protease/RNA-dependent RNA polymerase polypeptide precursor and capsid protein precursor, genes, complete cds; and unknown gene
7575	20537	33896	0.79	2.5E-01	U13982.1	NT	Mus musculus SKD1 (Skd1) gene, complete cds
7604	20565		1.13	2.5E-01	AF134119.1	NT	Arabidopsis thaliana DNA chromosome 4, contig fragment No. 18
7852	20788	34174	0.68	2.5E-01	AL161506.2	NT	Homo sapiens chromosome 21 segment HS21C082
7898	20841	34223	3.88	2.5E-01	AL163282.2	NT	767a03.x1 Scores_NSF_F8_9W_OT_PA_P_S1 Homo sapiens cDNA clone IMAGE:3525389 3'
8177	21147	34555	2.72	2.5E-01	BF109040.1	EST_HUMAN	601663391R2 NIH_MGC_58 Homo sapiens cDNA clone IMAGE:3826198 3'
8188	21158	34567	0.62	2.5E-01	BE960712.1	EST_HUMAN	601459238F1 NIH_MGC_66 Homo sapiens cDNA clone IMAGE:3862809 5'
8588	21536	34958	2.02	2.5E-01	BF038595.1	EST_HUMAN	E1B PROTEIN, SMALL T-ANTIGEN (E1B 19K)
8740	21708	35131	0.67	2.5E-01	P04492	SWISSPROT	y48407.1 Scores fetal liver spleen 1N1LS Homo sapiens cDNA clone IMAGE:202501 5'
8963	21949	35373	9.57	2.5E-01	HS3298.1	EST_HUMAN	Mouse testis-specific protein (TPX-1) gene, exon 10
9227	22193	35623	0.81	2.5E-01	M89626.1	NT	Homo sapiens matrix metalloproteinase MMP Rasi-1 gene, promoter region
9874	22827	36280	16.45	2.5E-01	U89851.2	NT	

Table 4

Single Exon Probes Expressed in Bone Marrow

Probe SEQ ID NO.	Exon SEQ ID NO.	ORF SEQ ID NO.	Expression Signal	Most Similar (Top) Hit BLAST E Value	Top Hit Accession No.	Top Hit Database Source	Top Hit Descriptor
9874	22827	36281	16.46	2.5E-01	U98851.2	NT	Homo sapiens matrix metalloproteinase MMP Rasi-1 gene, promoter region
9831	22814	36288	2.03	2.5E-01	AF085184.1	NT	Hordeum vulgare receptor-like kinase LRK10 gene, partial cds
9831	22814	36288	2.03	2.5E-01	AF085184.1	NT	Hordeum vulgare receptor-like kinase LRK10 gene, partial cds
10469	23380	36873	1.7	2.5E-01	AW581987.1	EST_HUMAN	RC3-ST0186-130100-015-e07 ST0186 Homo sapiens cDNA
10707	23629	37124	0.44	2.5E-01	11485852	NT	Porphyra purpurea chloroplast, complete genome
10809	23829	37342	1.58	2.5E-01	AW152246.1	EST_HUMAN	xg40c10.x1 NCI_CGAP_U11 Homo sapiens cDNA clone IMAGE:2630034 3' similar to contains Alu repetitive element; contains element MSR1 repetitive element;
10812	23832	37346	1.38	2.5E-01	X68481.1	NT	Mouse L1m1 LINE DNA
10892	23912	37427	0.45	2.5E-01	A0834721.1	EST_HUMAN	WP88611.x1 NCI_CGAP_Bm25 Homo sapiens cDNA clone IMAGE:2468876 3' similar to SW:CEBD_HUMAN P49716 CCAAT/ENHANCER BINDING PROTEIN DELTA;
10892	23912	37428	0.45	2.5E-01	A0834721.1	EST_HUMAN	WP88611.x1 NCI_CGAP_Bm25 Homo sapiens cDNA clone IMAGE:2468876 3' similar to SW:CEBD_HUMAN P49716 CCAAT/ENHANCER BINDING PROTEIN DELTA;
11413	24367	37892	4.47	2.5E-01	ID50914.1	NT	Human mRNA for KIAA0124 gene, partial cds
12120	24990	38591	2.34	2.5E-01	AE000711.1	NT	Aquifex aeolicus section 43 of 108 of the complete genome
12204	25046	38825	4.55	2.5E-01	AF200528.1	NT	Zea mays cellulose synthase-4 (Cesa-4) mRNA, complete cds
12230	25059		5.87	2.5E-01	AL161541.2	NT	Arabidopsis thaliana DNA chromosome 4, contig fragment No. 41
12691	25806	31627	1.54	2.5E-01	AF170072.1	NT	Spodoptera frugiperda CALNLC mRNA, complete cds
555	13825	26543	0.86	2.4E-01	A4938318.1	EST_HUMAN	en70d04.s1 Soares_NFL_T_GBC_S1 Homo sapiens cDNA clone IMAGE:1562023 3'
849	13905	26864	2.67	2.4E-01	BF576124.1	EST_HUMAN	802132442F1 NIH_MGC_61 Homo sapiens cDNA clone IMAGE:4271578 5'
1307	14343	27307	15.11	2.4E-01	AJ288880.1	NT	Homo sapiens KIAA0851 gene (partial), X13 gene and LZTFL1 gene
1307	14343	27308	15.11	2.4E-01	AJ288880.1	NT	Homo sapiens KIAA0851 gene (partial), X13 gene and LZTFL1 gene
1384	14418	27387	1.01	2.4E-01	Y17293.1	NT	Homo sapiens FLI-1 gene, partial
1888	14893		27.14	2.4E-01	AF267753.1	NT	Mesembryanthemum crystallinum putative potassium channel protein Mkt1p mRNA, complete cds
1916	14940	27836	1.21	2.4E-01	AF251708.1	NT	Zaocys diurnades fructose-1,6-bisphosphatase mRNA, complete cds
2148	15184	28180	0.97	2.4E-01	AF111168.2	NT	Homo sapiens serine palmitoyl transferase, subunit II gene, complete cds; and unknown genes
2178	15183		1.04	2.4E-01	P45384	SWISSPROT	IMMUNOGLOBULIN A1 PROTEASE PRECURSOR (IGA1 PROTEASE)
2271	15284	28310	2.23	2.4E-01	AE000680.1	NT	Aquifex aeolicus section 12 of 109 of the complete genome
2391	15399	28424	1.85	2.4E-01	BF002171.1	EST_HUMAN	7h23d04.x1 NCI_CGAP_C016 Homo sapiens cDNA clone IMAGE:3318807 3' similar to SW_FRSB_XENLA
2544	15548	28569	2.37	2.4E-01	Z96534.1	NT	D discoideum (Ax3-K) ponA gene
2772	15764	28785	2.11	2.4E-01	X71783.1	NT	S. pombe swi6 gene
2797	15789	28807	3.03	2.4E-01	AF030154.1	NT	Bovine adenovirus 3 complete genome
3149	16206		3.16	2.4E-01	U72726.1	NT	Oryza longistaminata receptor kinase-like protein, family member D, and retrofit (seg/pod) genes, complete cds

Page 87 of 546
Table 4

Single Exon Probes Expressed in Bone Marrow

Probe SEQ ID NO:	Exon SEQ ID NO:	ORF SEQ ID NO:	Expression Signal	Most Similar (Top) Hit BLAST E Value	Top Hit Accession No.	Top Hit Database Source	Top Hit Descriptor
3163	18218	28134	1.71	2.4E-01	X74209.1	NT	H. sapiens AGT gene, Ped fragment of intron 4
3675	18718	28632	1	2.4E-01	AF169783.1	NT	Podospora anserina HET-C protein (Het-C) gene, complete cds
3773	18815	28724	0.95	2.4E-01	AE000312.1	NT	Escherichia coli K-12 MG1655 section 202 of 400 of the complete genome
4092	17088		0.87	2.4E-01	D28980.1	NT	Rattus norvegicus mRNA for alphaB crystallin-related protein, complete cds
4974	17989		0.98	2.4E-01	AL101589.2	NT	Arabidopsis thaliana DNA chromosome 4, contig fragment No. 85
5089	18099	30875	0.95	2.4E-01	D00944.1	NT	Hepatitis C virus genomic RNA for polyprotein, complete cds
5537	18334	31574	0.77	2.4E-01	A925707.1	EST_HUMAN	wc33d05.x1 NCI_CGAP_Ges4 Homo sapiens cDNA clone IMAGE:2457129 3'
5537	18334	31575	0.77	2.4E-01	A925707.1	EST_HUMAN	wc33d05.x1 NCI_CGAP_Ges4 Homo sapiens cDNA clone IMAGE:2457129 3'
5532	18659	31605	0.89	2.4E-01	D50871.1	NT	Glycine max mRNA for mitotic cyclin b1-type, complete cds
5738	18832	32011	0.32	2.4E-01	AF091218.1	NT	Mus musculus Wm protein (Wm) gene, complete cds
5738	18832	32012	0.32	2.4E-01	AF091218.1	NT	Mus musculus Wm protein (Wm) gene, complete cds
5788	18858		0.65	2.4E-01	M83377.1	NT	Gallus gallus brain-derived neurotrophic factor (BDNF) gene, 5' end
5991	25649		1.03	2.4E-01	AJ138638.2	NT	Branchiostoma floridae mRNA for calmodulin 2 (calM2 gene)
5998	19082	32279	2.08	2.4E-01	BF692338.1	EST_HUMAN	754004.x1 NCI_CGAP_Br16 Homo sapiens cDNA clone IMAGE:3338503 3' similar to SW:SFR4_HUMAN
6098	19177	32395	2.07	2.4E-01	AF035546.1	NT	Q08170 SPLICING FACTOR, ARGININE/SERINE-RICH 4 ;contains element TAR1 repetitive element
6209	19283	32515	2.02	2.4E-01	7881801	NT	Drosophila melanogaster p38a MAP kinase gene, complete cds
6284	19337	32570	0.84	2.4E-01	AV733787.1	EST_HUMAN	Homo sapiens HSPC142 protein (HSPC142), mRNA
6523	19586	32844	0.69	2.4E-01	AA398672.1	EST_HUMAN	AV733787 cDNA Homo sapiens cDNA clone cdaADE11 5'
6885	19742	33018	1.72	2.4E-01	A1688989.1	EST_HUMAN	z70002.s1 Soares_testis_NHT Homo sapiens cDNA clone IMAGE:727683 3'
7247	19982	33279	0.54	2.4E-01	AF163983.1	NT	wc82c11.x1 NCI_CGAP_Pan1 Homo sapiens cDNA clone IMAGE:2323220 3' similar to gb.J03464
7593	20529	33888	7.83	2.4E-01	L43001.1	NT	PROCOLLAGEN ALPHA 2(I) CHAIN PRECURSOR (HUMAN);
7748	20701	34088	0.55	2.4E-01	N48732.1	EST_HUMAN	Mustela vison tyrosine aminotransferase gene, complete cds
7994	20833		0.51	2.4E-01	U05013.1	NT	Bos taurus guanylyl cyclase-activating protein 2 (guca2) mRNA, complete cds
7998	20835	34329	0.97	2.4E-01	AF228844.1	NT	My556c11.1 Soares_multiple_sclerosis_2Nbl-IMSP Homo sapiens cDNA clone IMAGE:277460 5'
8418	21385	34782	0.51	2.4E-01	X97252.1	NT	Rattus norvegicus Sprague-Dawley heme oxygenase-2 non-reducing isoform gene, complete cds
8416	21385	34783	0.51	2.4E-01	X97252.1	NT	Mus musculus path gene and promoter
8682	21860	35083	1.5	2.4E-01	AJ012585.1	NT	M. musculus path gene and promoter
8948	21912	35337	0.98	2.4E-01	BF242794.1	EST_HUMAN	Tetrahymena thermophila macronuclear gene encoding ribosomal protein L3, exons 1-2
9001	21887		0.57	2.4E-01	BF878275.1	EST_HUMAN	6018177679F1 NIH_MGC_65 Homo sapiens cDNA clone IMAGE:4106288 5'
9487	22451	35881	0.51	2.4E-01	AL139077.2	NT	602086188F1 NIH_MGC_83 Homo sapiens cDNA clone IMAGE:4250372 5'
							Campylobacter jejuni NCTC11168 complete genomes; segment 4/6

Table 4

Single Exon Probes Expressed in Bone Marrow

Probe SEQ ID NO:	Exon SEQ ID NO:	ORF SEQ ID NO:	Expression Signal	Most Similar (Top) Hit BLAST E Value	Top Hit Accession No.	Top Hit Database Source	Top Hit Descriptor
9487	22451	35892	0.51	2.4E-01	AL130077.2	NT	Campylobacter jejuni NCTC11168 complete genome; segment 4/8
9921	22742	36183	7.63	2.4E-01	AF033515.1	EST_HUMAN	wd43602.x1 Soares_NFL_T_QBC_S1 Homo sapiens cDNA clone IMAGE:23309008 3' similar to contains MER22J1 TAR1 repetitive element;
10082	22989	36457	0.57	2.4E-01	AF220067.1	NT	Drosophila melanogaster SKPB gene, complete cds
10082	22989	36458	0.57	2.4E-01	AF220067.1	NT	Drosophila melanogaster SKPB gene, complete cds
10805	23728	37227	1.59	2.4E-01	Q03682	SWISSPROT	COLLAGEN ALPHA 1(X) CHAIN PRECURSOR
11119	24078	37603	2.68	2.4E-01	AF146494.2	NT	Arabidopsis thaliana DNA chromosome 4, contig fragment No. 6
11185	24141	37675	1.77	2.4E-01	AF030189.1	NT	Mus musculus type 1 sigma receptor gene, complete cds
11534	24475		1.52	2.4E-01	Z21847.1	NT	P. asiatica mosaic virus genomic RNA
12098	24967	38584	4.86	2.4E-01	P06800	SWISSPROT	PROLINE RICH 33 KD EXTENSIN-RELATED PROTEIN PRECURSOR
12180	25015	38617	2.5	2.4E-01	AF217491.1	NT	Homo sapiens fragile 160 oxidoreductase (FOR) gene, exon 6
12284	25749		1.93	2.4E-01	AF004213.1	NT	Arabidopsis thaliana ethylene-insensitive3-like1 (EIL1) mRNA, complete cds
12361	25150		2.21	2.4E-01	AJ278191.1	NT	Mus musculus mRNA for putative mcr7 protein (mcr7 gene)
12575	25734		2.27	2.4E-01	V01507.1	NT	Gallus gallus gene coding for e-actin
12782	25946		1.31	2.4E-01	BF228876.1	EST_HUMAN	RC3-CT0413-100800-023-008 CT0413 Homo sapiens cDNA
13008	25559		8.49	2.4E-01	AL163281.2	NT	Homo sapiens chromosome 21 segment HS21C081
389	13484	26394	0.98	2.3E-01	S75898.1	NT	aromatase [Poecilia latipinna-zebra finches, ovary, mRNA, 3188 nt]
639	13704		5.85	2.3E-01	U89713.1	NT	Mycoplasma genitalium section 35 of 51 of the complete genome
668	13733	26658	21.34	2.3E-01	U87596.1	NT	Methanococcus jannaschii section 138 of 150 of the complete genome
934	13987	26937	4.19	2.3E-01	BE311893.1	EST_HUMAN	601142073F1 NIH_MGC_14 Homo sapiens cDNA clone IMAGE:3505818 5'
1511	14543	27514	1.33	2.3E-01	6577980	NT	Mus musculus vacuolar protein sorting 4b (Vps4b), mRNA
1568	14601		1.02	2.3E-01	U72837.2	NT	Yersinia pestis HmsH (hmsH), HmsF (hmsF), HmsR (hmsR), and HmsS (hmsS) genes, complete cds
1608	14640	27617	1.21	2.3E-01	AJ246480.1	NT	Brassica napus sig gene for S-locus glycoprotein, cultivar T2
1636	14688	27644	2.52	2.3E-01	Y10867.2	NT	Mus musculus cdh5 gene, exon 1, partial
2061	15079		1.33	2.3E-01	AJ235353.1	NT	Homo sapiens partial intron 3 of the wild type AF-4FEL gene
2452	15457	28476	2.76	2.3E-01	BE207718.1	EST_HUMAN	601175562F1 NIH_MGC_17 Homo sapiens cDNA clone IMAGE:3531015 5'
2560	15657	28876	1.27	2.3E-01	MT11318.1	NT	Human erythropoietin gene, complete cds
2835	14422	27391	2.34	2.3E-01	AB015033.1	NT	Martellabilla agrovocans gyrB gene for DNA gyrase subunit B, partial cds, strain:FO 14857
2874	16032	28955	1.26	2.3E-01	AA601379.1	EST_HUMAN	no16d08.s1 NCI_CGAP_Phet Homo sapiens cDNA clone IMAGE:1100843 3' similar to contains Alu repetitive element; contains element T1-IR repetitive element;
3100	16157		7.15	2.3E-01	R21732.1	EST_HUMAN	Yn21b07.s1 Soares placenta Nib24P Homo sapiens cDNA clone IMAGE:130357 3'
3393	18432	29360	1.26	2.3E-01	H09936.1	EST_HUMAN	yr97h10.1 Soares fetal liver spleen 1NPLS Homo sapiens cDNA clone IMAGE:213283 5'

Table 4

Single Exon Probes Expressed in Bone Marrow

Probe SEQ ID NO:	Exon SEQ ID NO:	ORF SEQ ID NO:	Expression Signal	Most Similar (Top) Hit BLAST E Value	Top Hit Accession No.	Top Hit Database Source	Top Hit Descriptor
3894	16894	28798	1.08	2.3E-01	S82821.1	NT	GSTA5=glutathione S-transferase Yc2 subunit (5' region, Intron 1) [rats, Morris hepatoma cell line, Genomic, 2212 nt, segment 1 of 3]
3958	16908		5.02	2.3E-01	7882133	NT	Homo sapiens KIAA0450 gene product (KIAA0450), mRNA
4212	17241		0.82	2.3E-01	J03287.1	NT	Rat atrial natriuretic factor (ANF) gene, 5' end
4377	17405	30285	0.94	2.3E-01	R82282.1	EST_HUMAN	y11761.1 Soares placenta Nb2H-IP Homo sapiens cDNA clone IMAGE:146017 5'
4428	17455		2.24	2.3E-01	L78789.1	NT	Mus musculus retin (Ret-1c) gene, promoter region
4479	17504	30392	1.07	2.3E-01	D90899.1	NT	Synechocystis sp. PCC6803 complete genome, 1/27, 1-133859
4515	17540	30426	1.9	2.3E-01	AF092535.1	NT	Homo sapiens mitogen-activated protein kinase p38delta (PRKM13) mRNA, complete cds
4584	17608	30502	7.42	2.3E-01	5031984	NT	Homo sapiens nuclear transport factor 2 (placental protein 15) (PP15) mRNA
5087	18097	30979	0.65	2.3E-01	AB032400.1	NT	Mus musculus tulip 1 mRNA, complete cds
5200	18209	31083	0.91	2.3E-01	U61328.1	NT	Human hereditary haemochromatosis region, histone 2A-like protein gene, hereditary haemochromatosis (HLA-H) gene, RefSeq gene, and sodium phosphate transporter (NP13) gene, complete cds
5377	18481	31356	2.19	2.3E-01	AB040845.1	NT	Homo sapiens mRNA for KIAA1512 protein, partial cds
5603	18603	31532	2.51	2.3E-01	BF058381.1	EST_HUMAN	7630606.x1 NC1 CGAP_Ov18 Homo sapiens cDNA clone IMAGE:3476868 3' similar to SW:GAG SMSAV
5608	18704	31801	4.59	2.3E-01	X86987.1	NT	P03330 GAG POLYPROTEIN [CONTAINS: CORE PROTEIN P16; INNER COAT PROTEIN P12; CORE SHELL PROTEIN P30; NUCLEOPROTEIN P10];
5733	18827		1.01	2.3E-01	L39112.1	NT	C.familialis rom1 gene
5845	18935	32119	0.81	2.3E-01	S60371.1	NT	Vitiforma cornu small subunit ribosomal RNA gene
6051	19132	32340	2.02	2.3E-01	A1708940.1	EST_HUMAN	23S rRNA [Leuconostoc carnosum, Genomic, 2666 nt]
							as27e12.x1 Barstead aorta HPLRB6 Homo sapiens cDNA clone IMAGE:2318446 3' similar to gb:X13238
6051	19132	32341	2.02	2.3E-01	A1708940.1	EST_HUMAN	CYTCHROME C OXIDASE POLYPEPTIDE VIC PRECURSOR (HUMAN);
							as27e12.x1 Barstead aorta HPLRB6 Homo sapiens cDNA clone IMAGE:2318446 3' similar to gb:X13238
6812	19886	33154	0.76	2.3E-01	AF198080.1	NT	CYTCHROME C OXIDASE POLYPEPTIDE VIC PRECURSOR (HUMAN);
							Oryctolagus cuniculus cytochrome oxidase subunit 1a (cox1a2) mRNA, complete cds; nuclear gene for mitochondrial product
							as42f12.x1 Barstead aorta HPLRB6 Homo sapiens cDNA clone IMAGE:2319887 3' similar to cortains Alu repetitive element;
7081	20083	33391	5.28	2.3E-01	A1718148.1	EST_HUMAN	Homo sapiens hypothalamic protein FLJ20345 (FLJ20345), mRNA
7318	20289	33632	0.89	2.3E-01	8823323	NT	Saccharomyces cerevisiae omega oxidase gene, complete cds
7505	20470	33830	0.79	2.3E-01	AF000227.1	NT	Glycine max resistance protein LM17 precursor RNA, partial cds
7646	20606	33972	2.65	2.3E-01	AF175389.1	NT	AV719681 GLC Homo sapiens cDNA clone GLCDB08 5'
7649	20609	33974	10.72	2.3E-01	AV719681.1	EST_HUMAN	AV719681 GLC Homo sapiens cDNA clone GLCDB08 5'
7649	20609	33976	10.72	2.3E-01	AV719681.1	EST_HUMAN	AV719681 GLC Homo sapiens cDNA clone GLCDB08 5'
7806	20810		3.39	2.3E-01	6754779	NT	Mus musculus myosin XV (Myo15), mRNA

Page 90 of 546
Table 4

Single Exon Probes Expressed in Bone Marrow

Probe SEQ ID NO:	Exon SEQ ID NO:	ORF SEQ ID NO:	Expression Signal	Most Similar (Top) Hit BLAST E Value	Top Hit Accession No.	Top Hit Database Source	Top Hit Descriptor
7871	20815	34183	1.36	2.3E-01	BE888071.1	EST_HUMAN	601511573F1 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:3912859 5'
8028	20985		2.69	2.3E-01	N80083.1	EST_HUMAN	zai2a08.r1 Soares fetal liver spleen 1NLS Homo sapiens cDNA clone IMAGE:282358 5'
8073	21010	34408	0.63	2.3E-01	11410821	NT	Homo sapiens protocadherin alpha cluster (LOC633660), mRNA
8073	21010	34409	0.63	2.3E-01	11410821	NT	Homo sapiens protocadherin alpha cluster (LOC633660), mRNA
8184	21154	34582	0.6	2.3E-01	AL161558.2	NT	Arabidopsis thaliana DNA chromosome 4, contig fragment No. 68
8328	21285		2.02	2.3E-01	M88931.1	NT	Oxytricha nova macronuclear telomere-binding protein alpha subunit (tel-alpha alanine version) gene, complete cds
8838	21805	36222	0.54	2.3E-01	U57999.1	NT	Mus musculus prosaposin (psap)(SGP-1) gene, complete cds
9125	22091	36519	0.46	2.3E-01	AW080541.1	EST_HUMAN	xc50e08.x1 NC1_OGAP_Bm35 Homo sapiens cDNA clone IMAGE:2591554 3'
9240	22206	36639	0.45	2.3E-01	AW984460.1	EST_HUMAN	EST376633 MAGE resequences, MAGH Homo sapiens cDNA
9498	22460	35900	0.59	2.3E-01	AA372164.1	EST_HUMAN	EST84061 Rhabdomyosarcoma Homo sapiens cDNA 5' end similar to DnaJ homolog (GB:X633368)
9498	22460	36901	0.69	2.3E-01	AA372164.1	EST_HUMAN	EST84061 Rhabdomyosarcoma Homo sapiens cDNA 5' end similar to DnaJ homolog (GB:X633368)
8940	22867	36329	0.63	2.3E-01	6879318	NT	Mus musculus phosphatidylinositol 3-kinase catalytic subunit delta (Pik3cd), mRNA
10086	23013	36486	0.76	2.3E-01	BE277890.1	EST_HUMAN	601120110F1 NIH_MGC_20 Homo sapiens cDNA clone IMAGE:2968739 5'
10141	23067	36543	0.69	2.3E-01	AW984460.1	EST_HUMAN	EST376533 MAGE resequences, MAGH Homo sapiens cDNA
10191	23116	36600	1.36	2.3E-01	X52124.1	NT	Haemophilus influenzae genes for HincII restriction-modification system (HincII methyltransferase (EC 2.1.1.72) and HincII endonuclease (EC 3.1.21.4))
10226	23151	36840	0.57	2.3E-01	AW364633.1	EST_HUMAN	PM2-DT0036-281289-001-404 DT0036 Homo sapiens cDNA
10264	23219	36702	2.8	2.3E-01	BE173060.1	EST_HUMAN	MR0-HT0559-240400-014-g11 HT0559 Homo sapiens cDNA
10351	23275	36749	2.33	2.3E-01	AJ283261.1	NT	Rhizobium leguminosarum partial genomic DNA for asopolysaccharide biosynthesis genes
10809	23730	37232	0.8	2.3E-01	AF201928.1	NT	Murine hepatitis virus strain 2, complete genome
10820	23741		6.12	2.3E-01	BF133577.1	EST_HUMAN	601646158R2 NIH_MGC_59 Homo sapiens cDNA clone IMAGE:4102082 3'
11355	24305	37832	1.49	2.3E-01	AF004833.1	NT	Mus musculus tissue factor pathway inhibitor (TFPI) mRNA, complete cds
11355	24305	37833	1.49	2.3E-01	AF004833.1	NT	Mus musculus tissue factor pathway inhibitor (TFPI) mRNA, complete cds
11522	24463	38015	1.69	2.3E-01	AJ250189.1	NT	Mus musculus partial mRNA for muscle protein 534 (mg534 gene)
11522	24463	38016	1.69	2.3E-01	AJ250189.1	NT	Mus musculus partial mRNA for muscle protein 534 (mg534 gene)
11679	24845	38222	2.61	2.3E-01	AE002167.2	NT	Chlamydia pneumoniae AR39, section 4 of 94 of the complete genome
12098	24968	38566	1.47	2.3E-01	AE004688.1	NT	Pseudomonas aeruginosa PA01, section 229 of 529 of the complete genome
12279	25098		5.42	2.3E-01	U45428.1	NT	Borrelia burgdorferi 2.9-6 locus, ORF-A-D genes, complete cds and REP+ gene, partial cds
12370	25156		6.49	2.3E-01	T27231.1	EST_HUMAN	HCOEST44 HT28M6 Homo sapiens cDNA clone HCOE44 5'
12403	25176		2.24	2.3E-01	AW863940.1	EST_HUMAN	PM4-SN0012-030400-001-b08 SN0012 Homo sapiens cDNA
12460	25892	31417	2.82	2.3E-01	AW303623.1	EST_HUMAN	xx21407.x1 Soares_NFL_T_GBC S1 Homo sapiens cDNA clone IMAGE:2813773 3' similar to TR-Q8Z175 Q8Z175 LYSYL OXIDASE-RELATED PROTEIN 2; contains PTR5.b2 TAR1 repetitive element;

Page 91 of 546
Table 4

Single Exon Probes Expressed in Bone Marrow

Probe SEQ ID NO:	Exon SEQ ID NO:	ORF SEQ ID NO:	Expression Signal	Most Similar (Top) Hit BLAST E Value	Top Hit Accession No.	Top Hit Database Source	Top Hit Descriptor
12494	25937	31312	11.07	2.3E-01	BE882484.1	EST_HUMAN	601507202F1 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:3908689 5'
12544	25287		2.36	2.3E-01	BF663319.1	EST_HUMAN	602144459F1 NIH_MGC_48 Homo sapiens cDNA clone IMAGE:4287719 5'
12591	25295		3.71	2.3E-01	AJ006519.1	NT	Rattus norvegicus mRNA for acid gated ion channel
12680	25348		2.12	2.3E-01	U48645.1	NT	Plasmodium waltl distal-less like protein PwDlx-3 (PwDlx-3) mRNA, complete cds
12628	25511		1.49	2.3E-01	BF475611.1	EST_HUMAN	ncs39h12.x1 Lupeid_sciatic_nerve Homo sapiens cDNA clone IMAGE:3395950 3' similar to contains element MEF38 repetitive element;
80	13206	26130	0.99	2.2E-01	AI052190.1	EST_HUMAN	oz14a10.x1 Soares_fetal_liver_spleen_1NFLS_S1 Homo sapiens cDNA clone IMAGE:1675230 3' similar to TR-Q13040 Q13040 ATP-BINDING CASSETTE PROTEIN;
1567	14800	27578	2.33	2.2E-01	AF187850.1	NT	Homo sapiens PPAR delta gene, promoter region
2100	15117	28139	2.2	2.2E-01	M34840.1	NT	Fresh-water sponge Enit1 alpha collagen (COLF-1) gene
2412	15419	28443	8.16	2.2E-01	BF677938.1	EST_HUMAN	602085608F1 NIH_MGC_83 Homo sapiens cDNA clone IMAGE:4249959 5'
2593	15594	28611	2.54	2.2E-01	BE618258.1	EST_HUMAN	601482829F1 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:3866190 5'
2593	15594	28612	2.54	2.2E-01	BE618258.1	EST_HUMAN	601482829F1 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:3866190 5'
2593	15594	28612	2.54	2.2E-01	BE618258.1	EST_HUMAN	PM2-HT0353-281289-003-e12 HT0353 Homo sapiens cDNA
2593	15594	28612	2.54	2.2E-01	BE618258.1	EST_HUMAN	PM2-HT0353-281289-003-e12 HT0353 Homo sapiens cDNA
2593	15594	28612	2.54	2.2E-01	BE618258.1	EST_HUMAN	Homo sapiens FRA3B common fragile region, diadenosine triphosphate hydrolase (FHT) gene, exon 5
2593	15594	28612	2.54	2.2E-01	BE618258.1	EST_HUMAN	Arabidopsis thaliana DNA chromosome 4, contig fragment No. 62
3403	16452		2.28	2.2E-01	AL161582.2	NT	Xiphophorus maculatus truncated Rex1 retrotransposon reverse transcriptase (RT) pseudogene
3827	16887		0.66	2.2E-01	AF155728.1	NT	Drosophila melanogaster UNC-119 (unc-119) gene, complete cds
4240	17269		1.2	2.2E-01	AF116102.1	NT	Mus musculus mixed lineage kinase 3 (Mlk3) and two pore domain K+ channel subunit (Kchld3) genes, complete cds
4247	17276	30158	5.88	2.2E-01	AF155142.1	NT	Mus musculus MAP kinase kinase kinase 1 (Mekk1) mRNA, complete cds
4282	17321	30200	2.33	2.2E-01	AF117340.1	NT	Mus musculus MAP kinase kinase kinase 1 (Mekk1) mRNA, complete cds
4292	17321	30201	2.33	2.2E-01	AF117340.1	NT	Human scRNA (BC200 beta) pseudogene
4383	17411	30294	1.21	2.2E-01	U01307.1	NT	Human scRNA (BC200 beta) pseudogene
4383	17411	30296	1.21	2.2E-01	U01307.1	NT	RC1-CT0249-141189-021-g04 CT0249 Homo sapiens cDNA
4456	17482	30370	23.17	2.2E-01	AW361098.1	EST_HUMAN	Human beta-cytoplasmic actin (ACTBP9) pseudogene
4663	17690		1.38	2.2E-01	D30804.1	NT	zfp87cd5.1 Stratagene INT neuron (#837233) Homo sapiens cDNA clone IMAGE:648998 5'
4663	17690	30773	1.67	2.2E-01	AA211216.1	EST_HUMAN	Mus musculus vinculin gene, exon 3
5082	18092		1.34	2.2E-01	L13298.1	NT	MR0-HT0067-201099-002-c10 HT0067 Homo sapiens cDNA
5160	18169	31048	1.22	2.2E-01	BE141035.1	EST_HUMAN	y42h09.1 Soares fetal liver spleen 1NFLS Homo sapiens cDNA clone IMAGE:208001 5' similar to gb:Z141118_maf1 CD59 GLYCOPROTEIN PRECURSOR (HUMAN);
5181	18190	31066	1.51	2.2E-01	H60648.1	EST_HUMAN	Homo sapiens chromosome 21 segment HS21C006
5263	18271		1.08	2.2E-01	AL163208.2	NT	

Table 4

Single Exon Probes Expressed in Bone Marrow

Probe SEQ ID NO:	Exon SEQ ID NO:	ORF SEQ ID NO:	Expression Signal	Most Similar (Top) Hit BLAST E Value	Top Hit Accession No.	Top Hit Database Source	Top Hit Descriptor
5938	18928	32112	2.18	2.2E-01	5803002	NT	Homo sapiens diaphanous (Drosophila, homolog) 2 (DIAPH2), transcript variant 156, mRNA
5949	18939		3.59	2.2E-01	D64000.1	NT	Synechocystis sp. PCC6803 complete genome, 1927, 2382729-2538999
6114	19192	32415	0.61	2.2E-01	U87087.1	NT	Gallus gallus T-box containing protein (Ch-Tbx1) mRNA, complete cds
6114	19192	32416	0.61	2.2E-01	U87087.1	NT	Gallus gallus T-box containing protein (Ch-Tbx1) mRNA, complete cds
6984	19917	33211	0.7	2.2E-01	AB038490.1	NT	Homo sapiens gene for fukutin, complete cds
6984	20189	33514	0.58	2.2E-01	AA490108.1	EST_HUMAN	ab02a09.s1 Stratagene fetal retina 937202 Homo sapiens cDNA clone IMAGE:838656 3'
6984	20189	33515	0.58	2.2E-01	AA490108.1	EST_HUMAN	ab02a09.s1 Stratagene fetal retina 937202 Homo sapiens cDNA clone IMAGE:838656 3'
7222	20244	33578	7.76	2.2E-01	AV756238.1	EST_HUMAN	AV756238 BM Homo sapiens cDNA clone BMFAHC06 5'
7336	20307	33650	1.43	2.2E-01	AF082738.1	NT	Streptococcus pyogenes phosphotidylglycerophosphate synthase (pgsA) and ABC transporter ATP-binding protein (stpA) genes, complete cds; and unknown genes
7336	20307	33651	1.43	2.2E-01	AF082738.1	NT	Streptococcus pyogenes phosphotidylglycerophosphate synthase (pgsA) and ABC transporter ATP-binding protein (stpA) genes, complete cds; and unknown genes
7509	20474	33634	2.11	2.2E-01	M24138.1	NT	Human glycophorin B gene, exon 4
7509	20474	33635	2.11	2.2E-01	M24138.1	NT	Human glycophorin B gene, exon 4
7728	20694	34048	0.59	2.2E-01	AE000035.2	NT	Mycoplasma pneumoniae M129 section 45 of 63 of the complete genome
7988	20905	34297	0.6	2.2E-01	AF287967.1	NT	Homo sapiens homeobox B7 (HOXB7) gene, partial cds; and homeobox B6 (HOXB6), homeobox B5 (HOXB5), homeobox B4 (HOXB4), and homeobox B3 (HOXB3) genes, complete cds
8000	20939	34332	0.5	2.2E-01	AB024553.1	NT	Bacillus halodurans DNA, complete and partial cds, strain C-125
8354	21323		2.51	2.2E-01	AF155143.1	NT	Mus musculus nm23-M1 gene, promoter region
8425	21394	34805	1	2.2E-01	Z49833.1	NT	E.coli sepA and sepB genes
8896	21892	35284	0.49	2.2E-01	AJ132918.1	NT	Pan troglodytes MeCP2 gene 3'UTR
9234	22200	35629	0.51	2.2E-01	L23312.1	NT	Mouse HD protein mRNA, complete cds
9234	22200	35630	0.51	2.2E-01	L23312.1	NT	Mouse HD protein mRNA, complete cds
9247	22213	35943	4.12	2.2E-01	AE001713.1	NT	Thermotoga maritima section 25 of 136 of the complete genome
9267	22293	35963	0.47	2.2E-01	U09984.1	NT	Mus musculus (CR) Swiss glyceroldehyde 3-phosphate dehydrogenase (Gapd-S) gene, complete cds
9376	22341		3.89	2.2E-01	AW856039.1	EST_HUMAN	PM3-CT0263-241268-009-507 CT0263 Homo sapiens cDNA
9470	22434	35972	1.4	2.2E-01	8393247	NT	Mus musculus deformed epidermal autoregulatory factor 1 (Drosophila) (Deaf1), mRNA
9553	22515	35986	1.4	2.2E-01	BF376354.1	EST_HUMAN	MR1-TN0045-110900-008-c02 TN0045 Homo sapiens cDNA
9644	22598	36037	1.3	2.2E-01	W02988.1	EST_HUMAN	za04f08.t1 Soares melanocyte 2N1H1M Homo sapiens cDNA clone IMAGE:291591 5'
9662	22619	36274	15.89	2.2E-01	P48634	SWISSPROT	LARGE PROLINE-RICH PROTEIN BAT2 (HLA-B-ASSOCIATED TRANSCRIPT 2)
9707	22660	36116	0.66	2.2E-01	AJ009839.1	NT	Xenopus laevis mRNA for kinesin-like protein 3 (xklp3)
9718	22746	36197	0.78	2.2E-01	7657428	NT	Mus musculus osteoblast specific factor 2 (OSF-2), mRNA
9731	22759	36212	4.38	2.2E-01	M89043.1	NT	Brachydanio rerio ependymin beta and gamma chains (Epd) gene, complete cds

Page 93 of 546
Table 4
Single Exon Probes Expressed in Bone Marrow

Probe SEQ ID NO:	Exon SEQ ID NO:	ORF SEQ ID NO:	Expression Signal	Most Similar (Top) Hit BLAST E Value	Top Hit Accession No.	Top Hit Database Source	Top Hit Descriptor
9978	22808	36371	0.5	2.2E-01	Q80880	SWISSPROT	CYC2LIC NUCLEOTIDE GATED CHANNEL, ROD PHOTORECEPTOR, ALPHA SUBUNIT (CNG CHANNEL 3) (CNG-3) (CNG3)
10174	23089	36579	3.74	2.2E-01	AF197941.1	NT	Funaria hygrometrica chloroplast-localized small heat shock protein (CP-sHSP21) mRNA, complete cds;
10313	23237	36719	1.78	2.2E-01	BF206607.1	EST_HUMAN	nuclear gene for chloroplast product
10537	23458	36856	1	2.2E-01	9825871	NT	601889724F1 NIH_MGC_19 Homo sapiens cDNA clone IMAGE:4100189 5'
10687	23619	37113	0.52	2.2E-01	T59472.1	EST_HUMAN	Human herpesvirus 6, complete genome
10687	23619	37114	0.52	2.2E-01	T69472.1	EST_HUMAN	y633008.t1 Stratiogene ovary (#6337217) Homo sapiens cDNA clone IMAGE:75855 5'
							y633008.t1 Stratiogene ovary (#6337217) Homo sapiens cDNA clone IMAGE:75855 5'
							Pseudomonas aeruginosa quorum-sensing ethanediol dehydrogenase (exaA) gene, partial cds; cytochrome c650 precursor (exaB), NAD+-dependent acetaldehyde dehydrogenase (exaC), and pyrroloquinoline quinone synthesis A (pqoA) genes, complete cds; and pyrroloquin
10735	23657	37150	0.51	2.2E-01	AF068264.1	NT	Mus musculus PHR1 (Phr1) gene, partial cds
10810	23731		0.68	2.2E-01	AF071001.1	NT	Helicobacter pylori, strain J99 section 123 of 132 of the complete genome
10856	23776	37274	0.67	2.2E-01	AE001562.1	NT	Helicobacter pylori, strain J99 section 123 of 132 of the complete genome
10856	23776	37275	0.67	2.2E-01	AE001562.1	NT	Drosophila 88C glue gene cluster
11751	24036	38215	3.77	2.2E-01	X01818.1	NT	Homo sapiens H-2K binding factor-2 (LOC51689), mRNA
11758	23944	37466	3.41	2.2E-01	7706216	NT	Homo sapiens chromosome Xq28 melanoma antigen family A2a (MAGEA2A), melanoma antigen family A12 (MAGEA12), melanoma antigen family A2b (MAGEA2B), melanoma antigen family A3 (MAGEA3), catractin (GALT), NAD(P)H dehydrogenase-like protein (NSDHL), and LP
12317	25950		2.32	2.2E-01	U82671.2	NT	Vitis vinifera cultivar Pinot Noir plasma membrane aquaporin (PIP1a) mRNA, complete cds
12406	25179		3.88	2.2E-01	AF188943.1	NT	RC1-CT0249-141199-021-g04 CT0249 Homo sapiens cDNA
12519	17482	30370	0.62	2.2E-01	AW361098.1	EST_HUMAN	h17b02.x1 NCI CGAP_GU1 Homo sapiens cDNA clone IMAGE:2972523 3'
12514	26248		1.7	2.2E-01	AW661922.1	EST_HUMAN	Photopus sungorus uncoupling protein 3 mRNA, partial cds
13111	25946		1.36	2.2E-01	AF271265.1	NT	nm31e11.s1 NCI CGAP_L1p2 Homo sapiens cDNA clone IMAGE:1061804
972	14024	26978	1.58	2.1E-01	AA669289.1	EST_HUMAN	Arabidopsis thaliana DNA chromosome 4, contig fragment No. 18
975	14026	26980	1.06	2.1E-01	AL161504.2	NT	Chlamydia muridarum, section 45 of 85 of the complete genome
1128	14169		2.38	2.1E-01	AE002314.2	NT	Mus musculus interferon (alpha and beta) receptor 2 (Ifnar2), mRNA
1203	14242	27169	0.88	2.1E-01	6754289	NT	Mus musculus interferon (alpha and beta) receptor 2 (Ifnar2), mRNA
1203	14242	27200	0.88	2.1E-01	6754289	NT	Mus musculus mas proto-oncogene and Igf2 gene for insulin-like growth factor type 2 and L41ps and Au78 pseudogenes
1512	14544	27515	1.05	2.1E-01	AJ248995.1	NT	ck73e02.s1 NCI CGAP_GC4 Homo sapiens cDNA clone IMAGE:1519610 3' similar to gb:K02765
1829	14953	27949	1.94	2.1E-01	AA906824.1	EST_HUMAN	COMPLEMENT C3 PRECURSOR (HUMAN);
2165	15161	28201	3.26	2.1E-01	BF695073.1	EST_HUMAN	602063128F1 NIH_MGC_81 Homo sapiens cDNA clone IMAGE:4247503 5'
2489	15885	28516	2.19	2.1E-01	H73968.1	EST_HUMAN	yu0407.s1 Scores fetal liver spleen 1N1FLS Homo sapiens cDNA clone IMAGE:232837 3'

Table 4

Single Exon Probes Expressed in Bone Marrow

Probe SEQ ID NO:	Exon SEQ ID NO:	ORF SEQ ID NO:	Expression Signal	Most Similar (Top) Hit BLAST E Value	Top Hit Accession No.	Top Hit Database Source	Top Hit Descriptor
2489	15895	28517	2.19	2.1E-01	H73968.1	EST_HUMAN	y04f07.s1 Soares fetal liver spleen 1N1FLS Homo sapiens cDNA clone IMAGE:232837 3'
2836	15894	28914	2	2.1E-01	6912445	NT	Homo sapiens potassium voltage-gated channel, subfamily H (eag-related), member 4 (KCNH4), mRNA
3461	16507	29428	0.92	2.1E-01	AA639482.1	EST_HUMAN	ng80b10.s1 NCI CGAP Cc8 Homo sapiens cDNA clone IMAGE:1158579 3'
3819	16939		6.5	2.1E-01	9838361	NT	Beta vulgaris mitochondrion, complete genome
4045	17083		1.16	2.1E-01	AE001793.1	NT	Thermotoga maritima section 105 of 136 of the complete genome
4084	17118	30013	1.48	2.1E-01	P11675	SWISSPROT	IMMEDIATE-EARLY PROTEIN IE180
4084	17118	30014	1.48	2.1E-01	P11675	SWISSPROT	IMMEDIATE-EARLY PROTEIN IE180
4403	17431		1.62	2.1E-01	AB033041.1	NT	Homo sapiens mRNA for KIAA1215 protein, partial cds
4607	17628	30520	1.75	2.1E-01	AB010273.1	NT	Homo sapiens pihsp47 gene, complete cds
5063	18073	30953	5.08	2.1E-01	D13567.1	NT	Lampetra japonica mRNA for alpha-2-macroglobulin, complete cds
5110	18120	30984	1.13	2.1E-01	Q01338	SWISSPROT	ALPHA-2A ADRENERGIC RECEPTOR (ALPHA-2A ADRENOCEPTOR) (ALPHA-2AAR)
5216	18225	31069	0.92	2.1E-01	AE001828.1	NT	Helicobacter pylori, strain J59 section 87 of 132 of the complete genome
5374	18479	31352	5.48	2.1E-01	BF672895.1	EST_HUMAN	602152001F1 NIH_MGC_81 Homo sapiens cDNA clone IMAGE:4283001 5'
7071	20093	33403	1.15	2.1E-01	AJ223392.1	NT	Dodo fragilis mitochondrial 16S rRNA gene, partial
7083	20017	33320	1.92	2.1E-01	U04642.1	NT	Human olfactory receptor (OR17-2) gene, partial cds
7636	20598	33959	0.78	2.1E-01	Q01956	SWISSPROT	VOLTAGE-GATED POTASSIUM CHANNEL PROTEIN KV3.3 (KSHIID)
7636	20598	33960	0.78	2.1E-01	Q01956	SWISSPROT	VOLTAGE-GATED POTASSIUM CHANNEL PROTEIN KV3.3 (KSHIID)
7648	20608		2.34	2.1E-01	AE000972.1	NT	Archaeoglobus fulgidus section 135 of 172 of the complete genome
7973	20912	34303	1.77	2.1E-01	AF000949.1	NT	Canis familiaris keratin (KRT9) gene, complete cds
8027	20984	34359	1.08	2.1E-01	AF068687.1	NT	Glycine max malate dehydrogenase (Mdh-2) gene, nuclear gene encoding mitochondrial protein, partial cds
8027	20984	34360	1.08	2.1E-01	AF068687.1	NT	Glycine max malate dehydrogenase (Mdh-2) gene, nuclear gene encoding mitochondrial protein, partial cds
8086	21022		0.51	2.1E-01	T87354.1	EST_HUMAN	y83501.r1 Soares fetal liver spleen 1N1FLS Homo sapiens cDNA clone IMAGE:114793 5'
8407	21378		1.04	2.1E-01	7305030	NT	Mus musculus erythrocyte protein band 4.1-like 3 (Epb4.1b), mRNA
8847	21814	35234	5.05	2.1E-01	U68399.1	NT	Haemophilus influenzae hmcD, putative haemocin processing protein (hmcC), putative ABC transporter (hmcB), putative haemocin structural protein (hmcA), and haemocin immunity protein (hmcI) genes, complete cds
9149	22115	35540	0.84	2.1E-01	AL040537.1	EST_HUMAN	DKFZp434H0614.J1 434 (synonym: htes3) Homo sapiens cDNA clone DKFZp434H0614 5'
9149	22115	35541	0.84	2.1E-01	AL040537.1	EST_HUMAN	DKFZp434H0614.J1 434 (synonym: htes3) Homo sapiens cDNA clone DKFZp434H0614 5'
9392	22397	35787	6.08	2.1E-01	Z35786.1	NT	S.cerevisiae chromosome II reading frame ORF YBL025W
9862	22798	36251	0.59	2.1E-01	N42536.1	EST_HUMAN	y71e10.r1 Soares melanocyte 2Nblm Homo sapiens cDNA clone IMAGE:270954 5'
9862	22798	36252	0.59	2.1E-01	N42536.1	EST_HUMAN	y71e10.r1 Soares melanocyte 2Nblm Homo sapiens cDNA clone IMAGE:270954 5'

Page 95 of 546
Table 4

Single Exon Probes Expressed in Bone Marrow

Probe SEQ ID NO:	Exon SEQ ID NO:	ORF SEQ ID NO:	Expression Signal	Most Similar (Top) BLAST E Value	Top Hit Accession No.	Top Hit Database Source	Top Hit Descriptor
9871	22824	38277	2.77	2.1E-01	X97378.1	NT	A.thaliana mRNA for AtRanBP1b protein
9878	22903	38387	1.28	2.1E-01	AB038629.1	NT	Homo sapiens p53R2 gene for ribonucleotide reductase, exon 6
10882	23614	37109	1.13	2.1E-01	Z97067.1	NT	Beta vulgaris mRNA for elongation factor 1-beta
10725	23847	37140	2.89	2.1E-01	P52824	SWISSPROT	DIACYLGLYCEROL KINASE, DELTA (DIGLYCERIDE KINASE) (DGK-DELTA) (DAG KINASE DELTA)
10731	23863	37146	0.93	2.1E-01	BF574254.1	EST_HUMAN	(80 KD DIACYLGLYCEROL KINASE) 602131427F1 NIH_MGC 81 Homo sapiens cDNA clone IMAGE:4270831 5'
11009	23974	37488	3.89	2.1E-01	AF294298.1	NT	Anolis lineatopus isolate NG NADH dehydrogenase subunit 2 (ND2) gene, complete cds; mitochondrial gene for mitochondrial product
11890	24771		2.15	2.1E-01	11036847	NT	Homo sapiens pancreatic polypeptide 2 (PPY2), mRNA
11905	24786	38378	1.81	2.1E-01	BE180422.1	EST_HUMAN	RC3-H170622-040500-013-b11 HT0622 Homo sapiens cDNA
12137	25502		1.38	2.1E-01	X57824.1	NT	Drosophila melanogaster ALA-E6 DNA, repeat region
12660	25339		1.94	2.1E-01	AF217490.1	NT	Homo sapiens fragile 16D octo reductase (FOR) gene, exons 8, 9, and partial cds
12880	25960		1.8	2.1E-01	L32888.1	NT	Human granulosa gene
12914	26493		2	2.1E-01	BE622149.1	EST_HUMAN	601440712F1 NIH_MGC 72 Homo sapiens cDNA clone IMAGE:3915675 5'
13045	25588	31882	1.62	2.1E-01	BE672330.1	EST_HUMAN	7a59e02 x1 NCI CGAP_G08 Homo sapiens cDNA clone IMAGE:3223034 3'
201	13302	26231	1.32	2.0E-01	AB017437.1	NT	Gallus gallus mRNA for avian, complete cds
535	13808		1.82	2.0E-01	7705601	NT	Homo sapiens CGA-18 protein (LOC51008), mRNA
700	13782	26694	1.25	2.0E-01	M77085.1	NT	O. cuniculus germline IgH heavy chain V-H pseudogene, allotype VH-2
811	13889	26818	1.96	2.0E-01	AF027865.1	NT	Mus musculus Major Histocompatibility Locus class II region
1013	14061	27012	0.72	2.0E-01	D68005.1	NT	Synechocystis sp. PCC6803 complete genome, 7/27, 781449-820815
1127	14170	27121	2.93	2.0E-01	AL163213.2	NT	Homo sapiens chromosome 21 segment HS21C013
1259	14294	27258	1.42	2.0E-01	AJ132685.5	NT	Homo sapiens rac1 gene
1311	14347	27313	2.04	2.0E-01	AW384637.1	EST_HUMAN	PM1-HT0422-291299-002-c08 HT0422 Homo sapiens cDNA
1454	14487		1.39	2.0E-01	AJ243857.1	NT	Plum pox virus strain M, complete genome, isolate PS
1483	14516	27490	15.27	2.0E-01	4503408	NT	Homo sapiens dystrobrevin, alpha (DTNA), mRNA
1555	14587	27569	2.13	2.0E-01	AB007974.1	NT	Homo sapiens mRNA, chromosome 1 specific transcript KIAA0505
1580	14592	27565	1.52	2.0E-01	AF260700.1	NT	Homo sapiens sodium/iodide symporter mRNA, partial cds
1708	14736	27718	1.38	2.0E-01	U22348.1	NT	Human bradykinin B1 receptor (bradyr1) gene, complete cds
1730	14760		2.09	2.0E-01	AF11170.3	NT	Homo sapiens 14q32 Jagged2 gene, complete cds; and unknown gene
1772	14801		3.8	2.0E-01	U87625.1	NT	Methanococcus jannaschii section 67 of 150 of the complete genome
1807	14931	27826	0.94	2.0E-01	BE871330.1	EST_HUMAN	601449441F1 NIH_MGC 65 Homo sapiens cDNA clone IMAGE:3853330 5'
1907	14931	27827	0.94	2.0E-01	BE871330.1	EST_HUMAN	601449441F1 NIH_MGC 65 Homo sapiens cDNA clone IMAGE:3853330 5'
2355	15384		1.88	2.0E-01	X82877.1	NT	H. sapiens Net-D-glucose cotransport regulator gene

Table 4

Single Exon Probes Expressed In Bone Marrow

Probe SEQ ID NO:	Exon SEQ ID NO:	ORF SEQ ID NO:	Expression Signal	Most Similar (Top) Hit BLAST E Value	Top Hit Accession No.	Top Hit Database Source	Top Hit Descriptor
3499	16546	29472	0.74	2.0E-01	P46607	SWISSPROT	HOMEOBOX PROTEIN GLABRA2 (HOMEOBOX-LEUCINE ZIPPER PROTEIN ATHB-10) (HD-ZIP PROTEIN ATHB-10)
3582	16827		0.82	2.0E-01	AW238005.1	EST_HUMAN	XP15022.x1 NC1_CGAP_JIN9 Homo sapiens cDNA clone IMAGE2740395 3' similar to contains element
3720	16763	29674	0.79	2.0E-01	P34641	SWISSPROT	MER21 repetitive element;
3855	16895	29789	1.14	2.0E-01	AL163204.2	NT	CED-11 PROTEIN
3977	17017	29831	0.7	2.0E-01	Z46906.1	NT	Homo sapiens chromosome 21 segment HS21C004
4597	17618		10.26	2.0E-01	BE826165.1	EST_HUMAN	Sus scrofa
5078	18088	30968	7.06	2.0E-01	8922080	NT	QV4-EN0032-190500-223-e03 EN0032 Homo sapiens cDNA
5144	18153	31033	0.97	2.0E-01	Y18216.1	NT	Homo sapiens hypothetical protein ASH1 (ASH1), mRNA
5298	18302	31163	0.93	2.0E-01	BE439491.1	EST_HUMAN	Homo sapiens putative pshHbD pseudogene for hair keratin, exons 1 to 9
5520	18619	31553	2.31	2.0E-01	X56900.1	NT	HTM1-122F1 HTM1 Homo sapiens cDNA
5831	18921	32104	1.89	2.0E-01	11432540	NT	Rat SOD-2 gene for manganese-containing superoxide dismutase
5940	19028	32220	0.69	2.0E-01	X91856.1	NT	Homo sapiens dual oxidase-like domains 2 (DUOX2), mRNA
6178	18253	32488	5.47	2.0E-01	U15300.1	NT	F. rubripes DNA encoding for vely-rRNA synthetase
6298	18370		0.79	2.0E-01	M75987.1	NT	Saccharomyces cerevisiae Hal5p (HAL5) mRNA, complete cds
6417	19484	32733	0.56	2.0E-01	P02467	SWISSPROT	Human hepatocyte growth factor gene, exon 1
6559	19829	32895	3.06	2.0E-01	X01033.1	NT	COLLAGEN ALPHA 2(I) CHAIN PRECURSOR
6679	19738	33012	4.25	2.0E-01	AW380885.1	EST_HUMAN	M. auratus mu class glutathione transferase gene
7512	20477	33838	1.27	2.0E-01	AF250371.1	NT	PM1-CT0247-141099-001-g06 CT0247 Homo sapiens cDNA
7679	20637	33899	0.72	2.0E-01	P54422	SWISSPROT	Mus musculus phosphofructokinase-1 C isozyme (Pfkfb) gene, exons 3 through 7
8064	20891	34389	0.61	2.0E-01	V00728.1	NT	GAMMA-GLUTAMYL TRANSPEPTIDASE PRECURSOR
8265	21254		0.17	2.0E-01	AF028028.1	NT	Mouse germ line gene coding for beta-globin (Y2)
8542	21510	34927	3.1	2.0E-01	X91151.1	NT	Andes virus strain O123133 glycoprotein G1 and G2 precursor, gene, partial cds
9074	22040		0.43	2.0E-01	BE562247.1	EST_HUMAN	M. musculus scp2 gene exon 14
9706	22659	38115	1.09	2.0E-01	U82511.1	NT	601344648F1 NIH_MGC_8 Homo sapiens cDNA clone IMAGE3677794 5'
9745	22888	38141	0.76	2.0E-01	U71122.1	NT	Dictyostelium discoideum slug cDNA19 protein (sc19) mRNA, partial cds
9914	22735		6.23	2.0E-01	AE001278.1	NT	Arabidopsis pyruvate decarboxylase-2 (Pdc2) gene, complete cds
10103	23029	38508	0.62	2.0E-01	P11420	SWISSPROT	Chlamydia trachomatis section 5 of 87 of the complete genome
10251	23176	38507	1.94	2.0E-01	AF146882.1	NT	DAUGHTERLESS PROTEIN
10401	23323	38807	1.94	2.0E-01	AF086907.1	NT	Homo sapiens filamin 2 (FLN2) mRNA, complete cds
10401	23323	38808	1.94	2.0E-01	AF086907.1	NT	Arabidopsis thaliana root gravitropism control protein (PIN2) gene, complete cds
10528	23450	38848	0.64	2.0E-01	AF157814.1	NT	Arabidopsis thaliana root gravitropism control protein (PIN2) gene, complete cds
							Homo sapiens cAMP specific phosphodiesterase (PDE4C) gene, exons 2 through 12

Table 4

Single Exon Probes Expressed in Bone Marrow

Probe SEQ ID NO:	Exon SEQ ID NO:	ORF SEQ ID NO:	Expression Signal	Most Similar (Top) Hit BLAST E Value	Top Hit Accession No.	Top Hit Database Source	Top Hit Descriptor
10528	23450	38849	0.84	2.0E-01	AF157814.1	NT	Homo sapiens cAMP specific phosphodiesterase (PDE4C) gene, exons 2 through 12
10576	23497		0.78	2.0E-01	X78388.1	NT	D.melanogaster DNA mobile element (hoppe)
10770	23691	37188	0.97	2.0E-01	X97121.1	NT	R.norvegicus mRNA for NTR2 receptor
11189	24145	37678	2.24	2.0E-01	D88088.1	NT	Salvelinus pluvius mRNA for transferrin, complete cds
11189	24145	37679	2.24	2.0E-01	D88088.1	NT	Salvelinus pluvius mRNA for transferrin, complete cds
12841	26325		1.42	2.0E-01	AF206837.2	NT	Pinophytes promelas liver glucose-6-phosphate-1-dehydrogenase mRNA, partial cds
12828	25779		1.66	2.0E-01	AF302773.1	NT	Homo sapiens nitric-Lm isoform (nitric) mRNA, complete cds
12837	25711	31609	1.34	2.0E-01	AW975297.1	EST_HUMAN	EST387405 MAGE resequences, MAGN Homo sapiens cDNA
12875	25508	31707	4.12	2.0E-01	A023592.1	EST_HUMAN	o60a10.s1 Soares testis_NHT Homo sapiens cDNA clone IMAGE:1843810 3'
12898	25483		2.98	2.0E-01	AF078164.2	NT	Homo sapiens Ku70-binding protein (KUB3) mRNA, partial cds
110	13221		9.9	1.9E-01	7549743	NT	Rattus norvegicus Aryl hydrocarbon receptor nuclear translocator 1 (Ahr1), mRNA
352	13441	26366	6.09	1.9E-01	AF004353.1	NT	Mus musculus pale ear (ep) gene, wild type allele, 3' region, partial cds
657	13723	26847	2.68	1.9E-01	U32581.2	NT	Homo sapiens lamda101a protein kinase C-interacting protein mRNA, complete cds
657	13723	26848	2.66	1.9E-01	U32581.2	NT	Homo sapiens lamda101a protein kinase C-interacting protein mRNA, complete cds
684	13730	26855	8.89	1.9E-01	BE070801.1	EST_HUMAN	RC3-BT0502-251199-011-401 BT0502 Homo sapiens cDNA
685	13730	26855	6.99	1.9E-01	BE070801.1	EST_HUMAN	RC3-BT0502-251199-011-401 BT0502 Homo sapiens cDNA
988	14039		1.19	1.9E-01	7305180	NT	Mus musculus interleukin 2 receptor, gamma chain (il2rg), mRNA
1108	14150	27101	9.83	1.9E-01	AA358813.1	EST_HUMAN	EST67784 Fetal lung II Homo sapiens cDNA 5' end
1372	14408	27376	2.9	1.9E-01	AF061282.1	NT	Sorghum bicolor 22 kDa keratin cluster
1430	14484		3.22	1.9E-01	AF184623.1	NT	Plasmodium vivax reticulocyte binding protein-2 (rbp-2) gene, complete cds
2390	15398	28423	3.89	1.9E-01	8922533	NT	Homo sapiens hypothetical protein FLJ10581 (FLJ10581), mRNA
2934	15922	28912	3.47	1.9E-01	U66088.1	NT	Sigmodon hispidus p53 gene, partial cds
2949	16006		4.89	1.9E-01	J00922.1	NT	Gallus gallus ovalbumin (Y) gene, complete cds
3408	16455	28378	4	1.9E-01	D13197.1	NT	Mouse gene for immunoglobulin diversity region D1
3492	16639	29484	5.36	1.9E-01	R16487.1	EST_HUMAN	Y42F10.11 Soares fetal liver spleen 1NFLS Homo sapiens cDNA clone IMAGE:128547 5'
3817	16857	29783	0.87	1.9E-01	AF284017.1	NT	Rattus norvegicus arylacetamide deacetylase gene, complete cds
3848	16888	29791	2.26	1.9E-01	P39788	SWISSPROT	PAIR-RULE PROTEIN ODD-PAIRED
4018	17057	28959	3.91	1.9E-01	AB006784.1	NT	Schistosoma japonicum pombe DNA for cytoplasmic dynein heavy chain, complete cds
4110	17144	30038	1.42	1.9E-01	AW754108.1	EST_HUMAN	CM3-CT0315-271199-045-b11 CT0315 Homo sapiens cDNA
4259	17288	30189	1.24	1.9E-01	BE5834943.1	EST_HUMAN	MR1-FN0010-280700-007-404 FN0010 Homo sapiens cDNA
4498	17521	30408	0.83	1.9E-01	AL161483.2	NT	Arabidopsis thaliana DNA chromosome 4, contig fragment No. 6
5047	18060		1.15	1.9E-01	AF223842.1	NT	Rattus norvegicus chemokine receptor CXCR3 mRNA, complete cds
5231	18239		1.01	1.9E-01	AK631198.1	EST_HUMAN	ISS9g12x1 NCI_QGAP_G08 Homo sapiens cDNA clone IMAGE:2238888 3' similar to gb:M21574 ALPHA PLATELET-DERIVED GROWTH FACTOR RECEPTOR PRECURSOR (HUMAN);

Page 98 of 546
Table 4

Single Exon Probes Expressed in Bone Marrow

Probe SEQ ID NO:	Exon SEQ ID NO:	ORF SEQ ID NO:	Expression Signal	Most Similar (Top) Hit BLAST E Value	Top Hit Accession No.	Top Hit Database Source	Top Hit Descriptor
5985	18780		5.2	1.9E-01	AW130149.1	EST_HUMAN	x22a07.x1 NCL CGAP_UH1 Homo sapiens cDNA clone IMAGE:2619444 3' similar to gb:M73779 RETINOIC ACID RECEPTOR ALPHA-1 (HUMAN);
5728	18822	32002	8.11	1.9E-01	AF127637.1	NT	Homo sapiens DNA polymerase epsilon catalytic subunit protein (POLE1) gene, exon 1a
5939	18025	32219	0.71	1.9E-01	AF081216.1	NT	Mus musculus Wm protein (Wm) gene, complete cds
5987	18072		2.15	1.9E-01	AU133116.1	EST_HUMAN	AU133116 NT2RP4 Homo sapiens cDNA clone NT2RP4001328 5'
6461	18528	32776	0.95	1.9E-01	AI762391.1	EST_HUMAN	w54h02.x1 NCL CGAP_Cot6 Homo sapiens cDNA clone IMAGE:2394089 3'
6527	18580	32850	0.88	1.9E-01	AW148452.1	EST_HUMAN	x14c08.x1 NCL CGAP_Ku8 Homo sapiens cDNA clone IMAGE:2618030 3' similar to gb:X03558 ATP SYNTHASE BETA CHAIN, MITOCHONDRIAL PRECURSOR (HUMAN);
7184	18395	31240	1.46	1.9E-01	R43212.1	EST_HUMAN	yg09a12.s1 Soares infant brain IN1B Homo sapiens cDNA clone IMAGE:31683 3' similar to contains MER13 repetitive element;
7193	20217	33547	0.95	1.9E-01	AF034920.1	NT	Homo sapiens tubby like protein 1 (TULP1) gene, exons 9-11
7193	20217	33548	0.95	1.9E-01	AF034920.1	NT	Homo sapiens tubby like protein 1 (TULP1) gene, exons 9-11
7474	20440	33798	0.65	1.9E-01	U73846.1	NT	Drosophila melanogaster testis-specific RNA-binding protein (bruno) mRNA, complete cds
7709	20688	34033	0.76	1.9E-01	U93688.1	NT	Staphylococcus aureus toxic shock syndrome toxin-1 (tsst), enterotoxin (ent), and integrase (int) genes, complete cds
7734	20689	34053	1.35	1.9E-01	U80922.1	NT	Arabidopsis thaliana serine/threonine protein phosphatase type one (TFPP8) gene, complete cds
7784	20737	34109	2.64	1.9E-01	AF072724.1	NT	Zea mays starch branching enzyme I (sbe1) gene, complete cds
8318	21287	34701	1.41	1.9E-01	AL161957.2	NT	Arabidopsis thaliana DNA chromosome 4, contig fragment No. 57
8034	22000	35421	12.98	1.9E-01	AB030024.1	NT	Homo sapiens mRNA for KIAA1188 protein, partial cds
9297	22263	35692	1.32	1.9E-01	M14568.1	NT	Marsupial cat beta-globin gene mRNA, partial cds
9297	22263	35693	1.32	1.9E-01	M14568.1	NT	Marsupial cat beta-globin gene mRNA, partial cds
10234	23159	36847	0.67	1.9E-01	AA912486.1	EST_HUMAN	db9g10.s1 NCL CGAP_PNS1 Homo sapiens cDNA clone IMAGE:1637608 3' similar to contains Alu repetitive element;
10802	23624	37019	0.72	1.9E-01	BE830353.1	EST_HUMAN	RC5-ET0082-060700-022-A02 ET0082 Homo sapiens cDNA
10802	23624	37020	0.72	1.9E-01	BE830353.1	EST_HUMAN	RC5-ET0082-060700-022-A02 ET0082 Homo sapiens cDNA
11106	24068	37588	1.87	1.9E-01	AF223391.1	NT	Homo sapiens calcium channel alpha1E subunit (CACNA1E) gene, exons 7-49, and partial cds, alternatively spliced
11631	24714	38298	1.47	1.9E-01	M22253.1	NT	Rattus norvegicus sodium channel 1 mRNA, complete cds
12033	24909	38504	3.54	1.9E-01	AJ243213.1	NT	Homo sapiens partial 6-HT4 receptor gene, exons 2 to 6
12049	24922	38519	1.53	1.9E-01	L07344.1	NT	Influenza A/Guangdong/2/97/72 nucleoprotein (seg 6) gene, 5' end
12875	25540		1.62	1.9E-01	AF036959.1	NT	Rattus norvegicus nuclear serine/threonine protein kinase mRNA, complete cds
34	13154	26055	2.78	1.9E-01	U79200.1	NT	Mus musculus p16Rip mRNA, complete cds
260	15637	26281	1.39	1.9E-01	AB022090.1	NT	Mus musculus Cctg gene for chaperonin containing TCP-1 gamma subunit, partial cds

Table 4

Single Exon Probes Expressed in Bone Marrow

Probe SEQ ID NO:	Exon SEQ ID NO:	ORF SEQ ID NO:	Expression Signal	Most Similar (Top) Hit BLAST E Value	Top Hit Accession No.	Top Hit Database Source	Top Hit Descriptor
370	13456	26386	1.91	1.8E-01	4502632	NT	Homo sapiens calcium channel, voltage-dependent, beta 2 subunit (CACNB2) mRNA, and translated products
983	14034	26886	0.89	1.8E-01	AI912212.1	EST_HUMAN	wf7102.x1 NCI_CGAP_Lu24 Homo sapiens cDNA clone IMAGE:2337051 3'
1093	14136	27087	1.05	1.8E-01	AF000580.1	NT	Dicotyledon discoidium plasmid Ddp5, complete genome
1293	14328	27290	9.43	1.8E-01	AL111789.1	NT	Yersinia pestis plasmid pCD1
1503	14536	27607	1.17	1.8E-01	6753947	NT	Mus musculus guanylate nucleotide binding protein 1 (Gbp1), mRNA
1503	14536	27608	1.17	1.8E-01	6753947	NT	Mus musculus guanylate nucleotide binding protein 1 (Gbp1), mRNA
1894	14889		1.37	1.8E-01	4505036	NT	Homo sapiens latent transforming growth factor beta binding protein 4 (LTBP4) mRNA
1882	14907		1.47	1.8E-01	AF733708.1	EST_HUMAN	qg22d10.x5 NCI_CGAP_Kid3 Homo sapiens cDNA clone IMAGE:1761811 3' similar to TR:O75936 O75936 GAMMA BUTYROBETAINE HYDROXYLASE :
1930	14954	27850	1.66	1.8E-01	AB051897.1	NT	Mus musculus Scy6, Scy6, Scy6, Scy6 genes for small inducible cytokine A6 precursor, small inducible cytokine A9 precursor, Scy6 pseudogene, small inducible cytokine A5 precursor, complete cds
2702	15698		3.94	1.8E-01	AW835728.1	EST_HUMAN	QV3-DT0018-081289-038-g04 DT0018 Homo sapiens cDNA
2908	15697		1.95	1.8E-01	AF184589.1	NT	Jonopsidium acule LEAFY protein (LEAFY2) gene, partial cds
2914	15672	28898	0.85	1.8E-01	AW182300.1	EST_HUMAN	Y41a03.x1 Soares_NFL_T_GBC_S1 Homo sapiens cDNA clone IMAGE:2659756 3'
3141	16198	29108	1.35	1.8E-01	AW995176.1	EST_HUMAN	QV0-BN0041-070300-147-c04 BN0041 Homo sapiens cDNA
3379	16429	29354	0.82	1.8E-01	BF183582.1	EST_HUMAN	601809723R1 NIH_MGC_18 Homo sapiens cDNA clone IMAGE:4040621 3'
3637	16680	29594	1.69	1.8E-01	H03369.1	EST_HUMAN	Y45e01.s1 Soares placenta Nb2HP Homo sapiens cDNA clone IMAGE:151704 3' similar to contains Alu repetitive element
3637	16680	29595	1.69	1.8E-01	H03369.1	EST_HUMAN	Y45e01.s1 Soares placenta Nb2HP Homo sapiens cDNA clone IMAGE:151704 3' similar to contains Alu repetitive element
4360	17387		0.75	1.8E-01	D37854.1	NT	Bovine NB25 mRNA for MHC class II (BdLA-DQB), complete cds
4596	17608	30503	6.42	1.8E-01	AL161586.2	NT	Arabidopsis thaliana DNA chromosome 4, contig fragment No. 58
4806	17823	30718	2.5	1.8E-01	AB051897.1	NT	Mus musculus Scy6, Scy6, Scy6, Scy6 genes for small inducible cytokine A6 precursor, small inducible cytokine A9 precursor, Scy6 pseudogene, small inducible cytokine A5 precursor, complete cds
5051	18063	30941	2.11	1.8E-01	X79794.1	NT	N. tabacum mRNA pNLA-35
5084	18094	30970	1.98	1.8E-01	AW814270.1	EST_HUMAN	MR3-ST0203-151299-112-g06 ST0203 Homo sapiens cDNA
5099	18109	30982	0.93	1.8E-01	AI792382.1	EST_HUMAN	en28g07.y5 Gessler Wilms tumor Homo sapiens cDNA clone IMAGE:1700028 5'
5131	18140	31018	2.98	1.8E-01	AF181258.1	NT	Mesocricetus auratus Na-taurocholate cotransporting polypeptide mRNA, partial cds
5214	18223		0.73	1.8E-01	U66150.1	NT	Loligo forbesi TTA repeat microsatellite region 1 for 4
5371	18476	31349	0.64	1.8E-01	BE082628.1	EST_HUMAN	RC8-BT0841-300300-011-H03 BT0841 Homo sapiens cDNA
5606	18982	32183	1.91	1.8E-01	AL161594.2	NT	Arabidopsis thaliana DNA chromosome 4, contig fragment No. 90

Table 4

Single Exon Probes Expressed in Bone Marrow

Probe SEQ ID NO.	Exon SEQ ID NO.	ORF SEQ ID NO.	Expression Signal	Most Similar (Top) Hit BLAST E Value	Top Hit Accession No.	Top Hit Database Source	Top Hit Descriptor
12301	25112	31840	1.88	1.8E-01	BF348823.1	EST_HUMAN	602018828F1 NCI_CGAP_Bm67 Homo sapiens cDNA clone IMAGE:4155318 5'
12752	25398		2.05	1.8E-01	Q80882	SWISSPROT	DNA TERMINAL PROTEIN (BELLETT PROTEIN) (PTP PROTEIN)
12868	25469		1.85	1.8E-01	R24494.1	EST_HUMAN	YH48H10.1 Soares placenta Nb2H1P Homo sapiens cDNA clone IMAGE:133027 5'
12907	25489		1.96	1.8E-01	Y11114.1	NT	E. dispar mRNA for hexokinase (hsk1)
579	13648	28561	1.77	1.7E-01	BE385194.1	EST_HUMAN	601274804F1 NIH_MGC_20 Homo sapiens cDNA clone IMAGE:3615768 5'
808	13884	28814	2.22	1.7E-01	X53330.1	NT	P. luminal histone gene cluster for core histones H2A, H2B, H3 and H4
862	14015		1.83	1.7E-01	P35816	SWISSPROT	NEUROFILAMENT TRIPLET L PROTEIN (NEUROFILAMENT LIGHT POLYPEPTIDE) (NFL)
1988	15019		3.18	1.7E-01	AF255051.1	NT	Homo sapiens BNIP3H (BNIP3H) gene, complete cds; nuclear gene for mitochondrial product
2871	15831	28848	2.3	1.7E-01	AF000718.1	NT	Vibrio cholerae hypoxanthine phosphoribosyltransferase (hpt) gene, partial cds, hemagglutinin/protease regulatory protein (hepR) gene, complete cds, and YRAL VIBCO gene, partial cds
2871	15931	28849	2.3	1.7E-01	AF000718.1	NT	Vibrio cholerae hypoxanthine phosphoribosyltransferase (hpt) gene, partial cds, hemagglutinin/protease regulatory protein (hepR) gene, complete cds, and YRAL VIBCO gene, partial cds
2838	15998	28917	1.69	1.7E-01	AA336909.1	EST_HUMAN	EST41651 Endometrial tumor Homo sapiens cDNA 5' end
3010	16068	28988	1.22	1.7E-01	AJ238736.1	NT	Naja naja atra chr-1 gene, exons 1-3
3010	16068	28989	1.22	1.7E-01	AJ238736.1	NT	Naja naja atra chr-1 gene, exons 1-3
3122	16179	29089	1.68	1.7E-01	AF081514.1	NT	Taxus canadensis geranylgeranyl diphosphate synthase mRNA, complete cds
3378	16427	29353	0.82	1.7E-01	N55763.1	EST_HUMAN	J2346F Human fetal heart, Lambda ZAP Express Homo sapiens cDNA clone J2348 5'
3462	16508	29429	1.48	1.7E-01	AJ269605.1	NT	Anabaena sp. ORF4 (partial), ORF3, ORF2, ORF1, adpA gene, adp8 gene, adpC gene, adpE gene and adpF gene
3959	16899	29814	5.89	1.7E-01	AJ263377.1	NT	Homo sapiens derivative 11 breakpoint fragment: partial intron 10 of the ALL-1/MLL/HRX gene fused to intron 5 of the AF-4/FEL gene
4691	17612		1.9	1.7E-01	X52936.1	NT	Schistosoma gregaria alpha repetitive DNA
4878	17893	30782	1.21	1.7E-01	AJ247635.1	EST_HUMAN	qh57e09.x1 Soares fetal liver spleen_1NFSL_S1 Homo sapiens cDNA clone IMAGE:1848808 3' similar to contains ORF.b1 ORF repetitive element;
5165	18174		0.82	1.7E-01	AF072725.1	NT	Zea mays starch branching enzyme IIb (se) gene, complete cds
5210	18219	31086	0.74	1.7E-01	BF030010.1	EST_HUMAN	601567258F1 NIH_MGC_58 Homo sapiens cDNA clone IMAGE:3827167 5'
5482	18582	31493	1.88	1.7E-01	AA470886.1	EST_HUMAN	ne13a02.s1 NCI_CGAP_C63 Homo sapiens cDNA clone IMAGE:881066 3' similar to gb:M17888 60S ACIDIC RIBOSOMAL PROTEIN P1 (HUMAN);
5482	18582	31494	1.88	1.7E-01	AA470886.1	EST_HUMAN	ne13a02.s1 NCI_CGAP_C63 Homo sapiens cDNA clone IMAGE:881066 3' similar to gb:M17888 60S ACIDIC RIBOSOMAL PROTEIN P1 (HUMAN);
5674	18768	31941	0.74	1.7E-01	U49589.1	NT	Brugia pahangi microfilarial sheath protein SHP3 (shp3) gene, complete cds
6463	19528	32777	12.89	1.7E-01	H72118.1	EST_HUMAN	ys02g06.s1 Soares fetal liver spleen 1NFSL Homo sapiens cDNA clone IMAGE:213688 3'

Table 4

Single Exon Probes Expressed in Bone Marrow

Probe SEQ ID NO:	Exon SEQ ID NO:	ORF SEQ ID NO:	Expression Signal	Most Similar (Top) Hit BLAST E Value	Top Hit Accession No.	Top Hit Database Source	Top Hit Descriptor
6628	19589	32848	0.9	1.7E-01	AB70978.1	EST_HUMAN	la20c11.x1 Soares_fetal_lung_NbHL19W Homo sapiens cDNA clone IMAGE:2045492 3'
6628	19589	32849	0.9	1.7E-01	AB70978.1	EST_HUMAN	la20c11.x1 Soares_fetal_lung_NbHL19W Homo sapiens cDNA clone IMAGE:2046492 3'
7031	18383	31250	0.78	1.7E-01	BE300288.1	EST_HUMAN	600944087T1 NIH_MGC_17 Homo sapiens cDNA clone IMAGE:2860248 3'
7033	20085		1.76	1.7E-01	AF028552.3	NT	Mesocricetus auratus oviductin precursor (OVI) gene, complete cds
7186	20220		0.79	1.7E-01	Z92910.1	NT	Homo sapiens HFE gene
7430	20397	33749	1.58	1.7E-01	AF000422.1	NT	Escherichia coli O157:H7 genomic DNA, Sakai-VT2 prophage inserted region
7515	20480	33841	8.55	1.7E-01	BE734178.1	EST_HUMAN	601569022F1 NIH_MGC_21 Homo sapiens cDNA clone IMAGE:3843984 5'
7721	20678	34043	1.42	1.7E-01	P16724	SWISSPROT	PROBABLE PROCESSING AND TRANSPORT PROTEIN UL66 (HFLF0 PROTEIN)
7739	26688	34057	0.71	1.7E-01	Q01955	SWISSPROT	COLLAGEN ALPHA 3(V) CHAIN PRECURSOR
8194	21164	34573	1.24	1.7E-01	AF000573.1	NT	Homo sapiens homogenitase 1,2-dioxygenase gene, complete cds
8296	21265	34676	0.82	1.7E-01	AF150689.1	NT	Pseudomonas putida long-chain-fatty-acid-CoA ligase (fadD) gene, complete cds
8820	21588	35004	7.37	1.7E-01	7708428	NT	Homo sapiens cleavage and polyadenylation specificity factor 3, 73kD subunit (CPSF3), mRNA
8620	21588	35005	7.37	1.7E-01	7708428	NT	Homo sapiens cleavage and polyadenylation specificity factor 3, 73kD subunit (CPSF3), mRNA
9045	22011	35433	0.6	1.7E-01	AW692873.1	EST_HUMAN	RC2-BN0032-120200-011-rt10 BN0032 Homo sapiens cDNA
9079	22045	35468	3.22	1.7E-01	D00384.1	NT	Rat (SHR strain) SX1 gene
9198	22162	35580	0.81	1.7E-01	AF217413.1	NT	Homo sapiens neuroigin 3 isoform gene, complete cds, alternatively spliced
9198	22162	35591	0.81	1.7E-01	AF217413.1	NT	Homo sapiens neuroigin 3 isoform gene, complete cds, alternatively spliced
9349	22314	35739	0.44	1.7E-01	RT7002.1	EST_HUMAN	y68g02.r1 Soares placenta Nb2-IP Homo sapiens cDNA clone IMAGE:144242 5'
9623	22486	35933	0.43	1.7E-01	BE253142.1	EST_HUMAN	601116872F1 NIH_MGC_16 Homo sapiens cDNA clone IMAGE:3357184 5'
9623	22486	35934	0.43	1.7E-01	BE253142.1	EST_HUMAN	601116872F1 NIH_MGC_16 Homo sapiens cDNA clone IMAGE:3357184 5'
9848	22875	36337	8.16	1.7E-01	AP001508.1	NT	Bacillus halodurans genomic DNA, section 2/14
10056	22983	36461	0.47	1.7E-01	AW977455.1	EST_HUMAN	EST388684 MAGe resequences, MAGO Homo sapiens cDNA
10056	22983	36462	0.47	1.7E-01	AW977455.1	EST_HUMAN	EST388684 MAGe resequences, MAGO Homo sapiens cDNA
10073	23000	36470	1.93	1.7E-01	U16288.1	NT	Human class IV alcohol dehydrogenase (ADH7) gene, exon 3
10148	23074	36548	0.59	1.7E-01	Z34508.1	NT	Human immunodeficiency virus type 1 (B7.05) env gene (partial)
10148	23074	36549	0.59	1.7E-01	Z34508.1	NT	Human immunodeficiency virus type 1 (B7.05) env gene (partial)
10167	23092	36570	0.7	1.7E-01	AJ251749.1	NT	Drosophila melanogaster mRNA for serine protease inhibitor (serpin-8), (sp8 gene)
10583	23515		2.43	1.7E-01	AL163284.2	NT	Homo sapiens chromosome 21 segment HS21C084
10759	23680	37176	1.24	1.7E-01	11427203	NT	Homo sapiens solute carrier family 7 (cationic amino acid transporter, y+ system), member 2 (SLC7A2), mRNA
10761	23682	37178	1.61	1.7E-01	AA827972.1	EST_HUMAN	nc60607.s1 NCL_CGAP_Co8 Homo sapiens cDNA clone IMAGE:1148282 3' similar to gb1.25081
10956	23986		0.42	1.7E-01	AL161542.2	NT	TRANSFORMING PROTEIN RHOC (HUMAN);
11040	24004	37530	8.17	1.7E-01	BE390835.1	EST_HUMAN	Arabidopsis thaliana DNA chromosome 4, contig fragment No. 42
							601286547F1 NIH_MGC_44 Homo sapiens cDNA clone IMAGE:3613258 5'

Table 4

Single Exon Probes Expressed in Bone Marrow

Probe SEQ ID NO:	Exon SEQ ID NO:	ORF SEQ ID NO:	Expression Signal	Most Similar (Top) Hit BLAST E Value	Top Hit Accession No.	Top Hit Database Source	Top Hit Descriptor
11157	24115	37841	2.61	1.7E-01	AA814617.1	EST_HUMAN	cd43a03.s1 NCL CGAP_CNS1 Homo sapiens cDNA clone IMAGE:1428924 3'
11447	24390	37933	8.03	1.7E-01	7106300	NT	Mus musculus adenomatosis polyposis coli binding protein Eb1 (Eb1), mRNA
11447	24390	37934	8.03	1.7E-01	7106300	NT	Mus musculus adenomatosis polyposis coli binding protein Eb1 (Eb1), mRNA
11703	24698	38245	1.6	1.7E-01	AA883375.1	EST_HUMAN	a45f08.s1 Soares_NFL_T_GBC_S1 Homo sapiens cDNA clone IMAGE:1460297 3'
12021	24898		1.68	1.7E-01	P15272	SWISSPROT	AMP NUCLEOSIDASE
12141	25005	38815	1.74	1.7E-01	11418157	NT	Homo sapiens calcium channel, voltage-dependent, alpha 1I subunit (CACNA1I), mRNA
12273	25933		1.54	1.7E-01	AL163278.2	NT	Homo sapiens chromosome 21 segment HS21C078
12437	25738	31618	1.55	1.7E-01	AA847421.1	EST_HUMAN	oe18h02.s1 NCL CGAP_O2 Homo sapiens cDNA clone IMAGE:1386281
							b69g05.s1 NCL CGAP_UH1 Homo sapiens cDNA clone IMAGE:2274872 3' similar to gb:M73779 RETINOIC ACID RECEPTOR ALPHA-1 (HUMAN);
12655	25738		1.69	1.7E-01	AB24404.1	EST_HUMAN	Human beta globin region on chromosome 11
12833	25450	31723	14.15	1.7E-01	U01317.1	NT	Homo sapiens melonate kinase gene, exon 6 and 7
126	13233	26162	2.02	1.6E-01	AF217532.1	NT	Y175f12.1 Soares placenta Nb2HP Homo sapiens cDNA clone IMAGE:135589 5'
680	15814	26370	1.56	1.6E-01	R31497.1	EST_HUMAN	nk28d12.s1 NCL CGAP_Cot11 Homo sapiens cDNA clone IMAGE:1014839 3'
1504	14537	27509	1.4	1.6E-01	AA648863.1	EST_HUMAN	Homo sapiens homeobox protein OTX2 gene, complete cds
1525	14557	27528	4.54	1.6E-01	AF296117.1	NT	AXONIN-1 PRECURSOR (AXONAL GLYCOPROTEIN TAG-1)
1941	14965	27982	1.79	1.6E-01	P22063	SWISSPROT	Crassostrea gigas RNA polymerase II largest subunit mRNA, partial cds
2001	15022		1	1.6E-01	U10334.1	NT	Homo sapiens cytochrome P450 3A4 (CYP3A4) gene, promoter region
2393	15892	28427	0.99	1.6E-01	X94232.1	NT	Populus trichocarpa cv. Trichobel ABI3 gene
2501	15504	28531	1.19	1.6E-01	AB037728.1	NT	Populus trichocarpa cv. Trichobel ABI3 gene
2902	15981	28881	10.42	1.6E-01	AF185589.1	NT	Archaeoglobus fulgidus section 145 of 172 of the complete genome
2902	15901	28982	10.42	1.6E-01	AF185589.1	NT	Vibrio cholerae chromosome II, section 70 of 93 of the complete chromosome
3649	16692	29607	1.2	1.6E-01	AJ003165.1	NT	Homo sapiens apelin gene, complete cds
3649	16692	29608	1.2	1.6E-01	AJ003165.1	NT	Homo sapiens apelin gene, complete cds
3788	16827	29734	0.78	1.6E-01	AE000982.1	NT	EST380877 MAGE resequences, MAGJ Homo sapiens cDNA
4025	17063		2.81	1.6E-01	AE004413.1	NT	Mus musculus chaperonin subunit 3 (gamma) (Cct3), mRNA
4356	17383	30266	10.43	1.6E-01	AF179880.1	NT	S.cerevisiae chromosome XI reading frame ORF YKR105c
4484	17509		3.1	1.6E-01	AW868601.1	EST_HUMAN	S.cerevisiae chromosome XI reading frame ORF YKR105c
4491	17516		4.57	1.6E-01	6753319	NT	z84f09.s1 Strabagene colon (#637204) Homo sapiens cDNA clone IMAGE:511361 3' similar to TR:E221955
4918	17833	30824	1.43	1.6E-01	Z28330.1	NT	E221955 38,855 BP SEGMENT OF CHROMOSOME XIV.;
4918	17833	30825	1.43	1.6E-01	Z28330.1	NT	Lycopodium obscurum Real fragment 2, satellite region
4982	17897	30885	4.36	1.6E-01	AA088343.1	EST_HUMAN	Lycopodium obscurum Real fragment 2, satellite region
5004	18018	30905	1.94	1.6E-01	AJ006358.1	NT	
5004	18018	30906	1.94	1.6E-01	AJ006358.1	NT	

Table 4

Single Exon Probes Expressed in Bone Marrow

Probe SEQ ID NO:	Exon SEQ ID NO:	ORF SEQ ID NO:	Expression Signal	Most Similar (Top) Hit BLAST E Value	Top Hit Accession No.	Top Hit Database Source	Top Hit Descriptor
5069	18079	30980	1.16	1.6E-01	BE018707.1	EST_HUMAN	b683h08.y1 NIH_MGC_10 Homo sapiens cDNA clone IMAGE:3046023 5' similar to gb:M61715
5461	18563	31475	0.87	1.6E-01	L40608.1	NT	TRYPTOPHANYL-TRNA SYNTHETASE (HUMAN); gb:X68857 M.musculus (MOUSE); Plasmodium falciparum (strain Dd2) variant-specific surface protein (var-1) gene, complete cds
5500	18598	31688	2.82	1.6E-01	AW197406.1	EST_HUMAN	xm43101.x1 NCI_CGAP_GC8 Homo sapiens cDNA clone IMAGE:2686969 5' similar to TR:O75984 O75984 HYPOTHETICAL 127.6 KD PROTEIN ;
5600	18696	31667	2.82	1.6E-01	AW197496.1	EST_HUMAN	xm43101.x1 NCI_CGAP_GC8 Homo sapiens cDNA clone IMAGE:2686969 3' similar to TR:O75984 O75984 HYPOTHETICAL 127.6 KD PROTEIN ;
5612	18708	31865	2.31	1.6E-01	AF034716.1	NT	Rattus norvegicus CCAAT/enhancer binding protein epsilon (cabepe) gene, complete cds
6144	19219	32449	0.9	1.6E-01	BE825903.1	EST_HUMAN	RC3-BN0034-310800-113-h01 BN0034 Homo sapiens cDNA
6385	19453	32697	0.57	1.6E-01	BF183394.1	EST_HUMAN	601809725R1 NIH_MGC_18 Homo sapiens cDNA clone IMAGE:4040335 3'
6385	19453	32698	0.57	1.6E-01	BF183394.1	EST_HUMAN	601809725R1 NIH_MGC_18 Homo sapiens cDNA clone IMAGE:4040335 3'
6568	19628	32883	1.99	1.6E-01	AL161688.2	NT	Arabidopsis thaliana DNA chromosome 4, contig fragment No. 84
6568	19628	32884	1.99	1.6E-01	AL161688.2	NT	Arabidopsis thaliana DNA chromosome 4, contig fragment No. 84
6850	20174	33498	0.56	1.6E-01	AA388047.1	EST_HUMAN	289404.J1 Soares testis_NHT Homo sapiens cDNA clone IMAGE:728511 5'
6869	20192	33519	0.66	1.6E-01	AB046788.1	NT	Homo sapiens mRNA for KIAA1568 protein, partial cds
7158	18388	31232	4.63	1.6E-01	AW291215.1	EST_HUMAN	UHH-B12-ag1-b-08-Q-U1.s1 NCI_CGAP_Sub4 Homo sapiens cDNA clone IMAGE:2724418 3'
7518	20483	33844	0.61	1.6E-01	Z49632.1	NT	S.cerevisiae chromosome X reading frame ORF YJR132w
8056	20983	34390	1.69	1.6E-01	AW248359.1	EST_HUMAN	2822248.Sprime NIH_MGC_7 Homo sapiens cDNA clone IMAGE:2822248 5'
8060	21026	34425	0.57	1.6E-01	6753237	NT	Mus musculus Ccr-2>-dependent activator protein for secretion (Cadps), mRNA
8102	21038		1.15	1.6E-01	AU136525.1	EST_HUMAN	AU136525 PLACE1 Homo sapiens cDNA clone PLACE1004488 5'
8201	21171	34582	1.28	1.6E-01	L48349.1	NT	Gorilla gorilla androgen receptor gene, partial exon
8359	21328		0.51	1.6E-01	BE244087.1	EST_HUMAN	TCBAP1E0807 Pediatric pro-B cell acute lymphoblastic leukemia Baylor-HGSC project=TCBA Homo sapiens cDNA clone TCBAP0607
8456	21425	34841	0.66	1.6E-01	U38243.1	NT	Bacteroides vulgatus beta-lactamase (cbva) gene, complete cds and mobilization protein (mobA) gene, complete cds
8979	21945	35369	0.85	1.6E-01	Z99119.1	NT	Bacillus subtilis complete genome (section 16 of 21); from 2997771 to 3213410
9178	22144	35571	0.71	1.6E-01	R13673.1	EST_HUMAN	y607h08.J1 Soares infant brain 1N1B Homo sapiens cDNA clone IMAGE:28873 5'
9285	22251		0.63	1.6E-01	L36861.1	NT	Homo sapiens guanylate cyclase activating protein (GCAP) gene exons 1-4, complete cds
9322	22287	35717	1.9	1.6E-01	Z49501.1	NT	S.cerevisiae chromosome X reading frame ORF YJR001w
9468	22430		0.63	1.6E-01	AF111167.2	NT	Homo sapiens Jun dimerization protein gene, partial cds; cfos gene, complete cds; and unknown gene
10009	22636		2.05	1.6E-01	BF375171.1	EST_HUMAN	RC3-ST0200-041189-011-h01 ST0200 Homo sapiens cDNA
10012	22639	36404	2	1.6E-01	Z49501.1	NT	S.cerevisiae chromosome X reading frame ORF YJR001w
10048	22875		0.86	1.6E-01	BE155684.1	EST_HUMAN	PM2-HT0353-270100-004-111 HT0353 Homo sapiens cDNA

Table 4

Single Exon Probes Expressed in Bone Marrow

Probe SEQ ID NO:	Exon SEQ ID NO:	ORF SEQ ID NO:	Expression Signal	Most Similar (Top) Hit BLAST E Value	Top Hit Accession No.	Top Hit Database Source	Top Hit Descriptor
11017	23982	37509	2.59	1.6E-01	AW850853.1	EST_HUMAN	IL3-CT0220-11188-028-G01 GT0220 Homo sapiens cDNA
11122	24082	37607	1.82	1.6E-01	Z28073.1	NT	S.cerevisiae chromosome XI reading frame ORF YKL073W
11122	24082	37608	1.82	1.6E-01	Z28073.1	NT	S.cerevisiae chromosome XI reading frame ORF YKL073W
11344	24294	37820	1.5	1.6E-01	BE258648.1	EST_HUMAN	601145780F1 NIH_MGC_19 Homo sapiens cDNA clone IMAGE:3161183 5'
11451	24394		3.84	1.6E-01	AF108064.1	NT	Plasmodium falciparum calcium-dependent protein kinase-3 (cdpk3) gene, complete cds
11740	24625	38204	6.59	1.6E-01	6671552	NT	Mus musculus adaptor-related protein complex AP-1, beta 1 subunit (Apt1b1), mRNA
12155	25006		1.75	1.6E-01	6679468	NT	Mus musculus protein kinase, cGMP-dependent, type II (Prkg2), mRNA
12275	25095	38179	6.76	1.6E-01	AV718685.1	EST_HUMAN	AV718685 GLC Homo sapiens cDNA clone G1CEMF07 5'
12608	25307		1.02	1.6E-01	AW839711.1	EST_HUMAN	RC1-LT0074-120200-014401_1 LT0074 Homo sapiens cDNA
12697	25721		22.15	1.6E-01	AB045310.1	NT	Cucurbita sativus K3 mRNA for anti-leucine synthase, complete cds
12856	25463		2.73	1.6E-01	AK024498.1	NT	Homo sapiens mRNA for FLJ010104 protein, partial cds
12945	25524		2.56	1.6E-01	AF287344.1	NT	Fuchsia hybrid cultivar Qiu 94208 ribosomal protein S10 gene, partial cds; nuclear gene for mitochondrial product
12971	25637	31716	2.24	1.6E-01	9506522	NT	Rattus norvegicus chondroitin sulfate proteoglycan 5 (neuroglycan C) (Cspg5), mRNA
12978	25543		1.63	1.6E-01	BE267894.1	EST_HUMAN	601126459F1 NIH_MGC_8 Homo sapiens cDNA clone IMAGE:3345038 5'
248	13345	26270	1.73	1.5E-01	BE710087.1	EST_HUMAN	IL3-HT0619-040700-197-E05 HT0619 Homo sapiens cDNA
248	13346	26271	1.73	1.5E-01	BE710087.1	EST_HUMAN	IL3-HT0619-040700-197-E05 HT0619 Homo sapiens cDNA
589	15813		2.4	1.5E-01	AV711698.1	EST_HUMAN	AV711698 DCA Homo sapiens cDNA clone DCAADH08 5'
783	13943	26768	1.51	1.5E-01	AL163284.2	NT	Homo sapiens chromosome 21 segment HS21C084
1094	14138	27089	1.01	1.5E-01	AJ009735.1	NT	Cyprinus carpio mRNA for EGGS22 myosin heavy chain, 3'UTR
1099	14143	27093	2.55	1.5E-01	AJ251885.1	NT	Homo sapiens partial SLC22A2 gene for organic cation transporter (OCT2), exon 1
1115	14159		1.57	1.5E-01	L36125.1	NT	Rattus norvegicus insulin-responsive glucose transporter (GLUT4) gene, 5' end
1280	14315	27278	3.58	1.5E-01	D28535.1	NT	Human gene for dihydrolipoamide succinyltransferase, complete cds (exon 1-15)
1280	14315	27277	3.58	1.5E-01	D28535.1	NT	Human gene for dihydrolipoamide succinyltransferase, complete cds (exon 1-15)
1478	14511	27487	2.54	1.5E-01	AF117340.1	NT	Mus musculus MAP kinase kinase kinase 1 (Mekk1) mRNA, complete cds
1823	14947	27943	1.65	1.5E-01	AW444451.1	EST_HUMAN	UHL-B13-atb-b-09-0-UL1.1 NCJ_CGAP_Sub5 Homo sapiens cDNA clone IMAGE:2733641 3'
2723	15717	28736	1.47	1.5E-01	BF686381.1	EST_HUMAN	602083238F1 NIH_MGC_81 Homo sapiens cDNA clone IMAGE:4247537 5'
							xw56a02.x2 NCJ_CGAP_Pan1 Homo sapiens cDNA clone IMAGE:2831078 3' similar to gb:U56072_mn1
2824	15982		1.1	1.5E-01	AW572516.1	EST_HUMAN	THYROID HORMONE RECEPTOR ALPHA-1 (HUMAN);
3048	16105	29019	0.78	1.5E-01	M81441.1	NT	Bos taurus factor V variant 2 (factor V) mRNA, complete cds
3363	16413	29338	5.55	1.5E-01	AA935048.1	EST_HUMAN	cc88d06.s1 NCJ_CGAP_GC4 Homo sapiens cDNA clone IMAGE:1671337 3' similar to gb:M11433
3381	16430	29367	0.74	1.5E-01	Z23104.1	NT	RETINOL-BINDING PROTEIN I, CELLULAR (HUMAN);
3381	16430	29358	0.74	1.5E-01	Z23104.1	NT	L.stagnalis mRNA for G protein-coupled receptor

Table 4

Single Exon Probes Expressed in Bone Marrow

Probe SEQ ID NO:	Exon SEQ ID NO:	ORF SEQ ID NO:	Expression Signal	Most Similar (Top) Hit BLAST E Value	Top Hit Accession No.	Top Hit Database Source	Top Hit Descriptor
3440	18487	29405	0.97	1.5E-01	AW612237.1	EST_HUMAN	h12802.x1 NCI_CGAP_Lu24 Homo sapiens cDNA clone IMAGE:2956539 3' similar to contains element MER16 repetitive element;
3768	18810	29719	2.22	1.5E-01	U09884.1	NT	Mus musculus ICR/Swiss glyceroldehyde 3-phosphate dehydrogenase (Gapd-S) gene, complete cds Homo sapiens pyruvate dehydrogenase kinase, isoenzyme 1 (PDK1), nuclear gene encoding mitochondrial protein, mRNA
3782	18823	29731	0.8	1.5E-01	7108358	NT	XYNA; Thermotoga bacterium; xynA; 4182 base-pairs
3785	18835	29740	0.66	1.5E-01	M87882.1	NT	h10100.x1 Soares_NFL_T_GBC_S1 Homo sapiens cDNA clone IMAGE:2881411 3'
3881	18920	29829	2.26	1.5E-01	AW665983.1	EST_HUMAN	Populus trichocarpa cv. Trichobel ABIS gene
3896	18936	29846	0.71	1.5E-01	AJ003165.1	NT	Populus trichocarpa cv. Trichobel ABIS gene
3898	18938	29847	0.71	1.5E-01	AJ003165.1	NT	Populus trichocarpa cv. Trichobel ABIS gene
4080	17115	30011	2.63	1.5E-01	AW368859.1	EST_HUMAN	RC2-HTT0149-191098-012-c09 HT0149 Homo sapiens cDNA
4213	17242	30127	9.83	1.5E-01	AL163284.2	NT	Homo sapiens chromosome 21 segment HS21C084
4513	17538	30422	0.91	1.5E-01	BE791253.1	EST_HUMAN	601583938F1 NIH_MGC 7 Homo sapiens cDNA clone IMAGE:3938345 5'
4513	17538	30423	0.91	1.5E-01	BE791253.1	EST_HUMAN	601583938F1 NIH_MGC 7 Homo sapiens cDNA clone IMAGE:3938345 5'
4768	17776	30872	1.88	1.5E-01	BF687685.1	EST_HUMAN	602067192F1 NIH_MGC 57 Homo sapiens cDNA clone IMAGE:4066223 5'
4783	15717	28735	2.88	1.5E-01	BF685391.1	EST_HUMAN	602083289F1 NIH_MGC 81 Homo sapiens cDNA clone IMAGE:4247637 5'
4830	17847	30747	1	1.5E-01	BE173798.1	EST_HUMAN	CMD-HT0565-280200-245-b10 HT0565 Homo sapiens cDNA
4830	17847	30748	1	1.5E-01	BE173798.1	EST_HUMAN	CMD-HT0565-280200-245-b10 HT0565 Homo sapiens cDNA
5034	18048	30828	1.2	1.5E-01	AL161580.2	NT	Arabidopsis thaliana DNA chromosome 4, contig fragment No. 60
5134	18143	31023	1.07	1.5E-01	5579451	NT	Homo sapiens calbindin 1, (28kD) (CALB1), mRNA
5328	18434	31188	2.46	1.5E-01	P07998	SWISSPROT	THROMBOSPONDIN 1 PRECURSOR
5357	18482	31331	1.03	1.5E-01	AF256882.1	NT	Callinectes sapidus MHC class II beta chain (hclbeta) gene, complete cds
5401	18504		6.6	1.5E-01	P15198	SWISSPROT	SEX HORMONE-BINDING GLOBULIN PRECURSOR (SHBG) (SEX STEROID-BINDING PROTEIN) (SBP)
5616	18712	31870	4.35	1.5E-01	AW650754.1	EST_HUMAN	(IL3-CT0219-160200-084-F10 CT0219 Homo sapiens cDNA
5659	18755	31822	7.17	1.5E-01	U65018.1	NT	Mus musculus transforming growth factor alpha (TGFA) mRNA, complete cds
5659	18755	31823	7.17	1.5E-01	U65018.1	NT	Mus musculus transforming growth factor alpha (TGFA) mRNA, complete cds
6012	18066	32286	0.79	1.5E-01	4506810	NT	Homo sapiens sodium channel, voltage-gated, type VI, alpha polypeptide (SCN6A) mRNA
6120	19198	32422	1.74	1.5E-01	6753659	NT	Mus musculus DNA methyltransferase 2 (Dnm2), mRNA
6120	19198	32423	1.74	1.5E-01	6753659	NT	Mus musculus DNA methyltransferase 2 (Dnm2), mRNA
6161	18236	32467	1.98	1.5E-01	AJ276505.1	NT	Mus musculus genomic fragment, 279 Kb, chromosome 7
6319	18390	32630	3.25	1.5E-01	BE727698.1	EST_HUMAN	601584322F1 NIH_MGC 20 Homo sapiens cDNA clone IMAGE:3833981 5'
6375	18443		1.86	1.5E-01	4508396	NT	Homo sapiens RAD54 (S.cerevisiae)-like (RAD54L) mRNA

Table 4

Single Exon Probes Expressed in Bone Marrow

Probe SEQ ID NO:	Exon SEQ ID NO:	ORF SEQ ID NO:	Expression Signal	Most Similar (Top) Hit BLAST E Value	Top Hit Accession No.	Top Hit Database Source	Top Hit Descriptor
6478	19543	32791	1.7	1.5E-01	AF134907.1	NT	Influenza B virus (BNanchang/480/94) NB protein gene, complete cds; and neuraminidase gene, partial cds
6651	25684	32985	1.8	1.5E-01	AE001039.1	NT	Archaeoglobus fulgidus section 68 of 172 of the complete genome
6681	19738	33013	4.83	1.5E-01	11417238	NT	Homo sapiens chromosome 5 open reading frame 3 (C6ORF3), mRNA
6682	19749	33026	1.8	1.5E-01	P48508	SWISSPROT	GLUTAMATE-CYSTEINE LIGASE REGULATORY SUBUNIT (GAMMA-GLUTAMYL-CYSTEINE SYNTHETASE) (GAMMA-ECS) (GCS LIGHT CHAIN)
6740	19795	33075	2.35	1.5E-01	Q28462	SWISSPROT	AMELOGENIN
6842	19895	33188	0.85	1.5E-01	AA714780.1	EST_HUMAN	mw30d10.s1 NCL_CGAP_GCB0 Homo sapiens cDNA clone IMAGE:1241971 3'
6871	19924	33220	1.8	1.5E-01	P30143	SWISSPROT	HYPOTHETICAL 51.7 KD PROTEIN IN THRC-TALB INTERGENIC REGION (ORF8)
7171	19402	31247	5.15	1.5E-01	AW970285.1	EST_HUMAN	EST382376 MAGE resequencing, MAGK Homo sapiens cDNA
7214	25678		0.71	1.5E-01	AA811545.1	EST_HUMAN	cb73f02.s1 NCL_CGAP_GCB1 Homo sapiens cDNA clone IMAGE:1337019 3' similar to contains element LTR2 repetitive element;
7424	20391		1.91	1.5E-01	AF210842.1	NT	Homo sapiens HARP (HARP) gene, exon 17 and complete cds
7621	20581	33945	1.88	1.5E-01	AF973157.1	EST_HUMAN	wr32cd08.x1 NCL_CGAP_UH1 Homo sapiens cDNA clone IMAGE:2481310 3'
7848	20783	34168	0.98	1.5E-01	AF29073.1	NT	Bos taurus Niemann-Pick type C1 disease protein (NPC1) mRNA, complete cds
7848	20783	34170	0.96	1.5E-01	AF29073.1	NT	Bos taurus Niemann-Pick type C1 disease protein (NPC1) mRNA, complete cds
7857	20802	34177	1.84	1.5E-01	AW500811.1	EST_HUMAN	UI-HF-BNO-ekk-d-05-0-UI.1 NIH_MGC_50 Homo sapiens cDNA clone IMAGE:3077409 5'
7857	20802	34178	1.84	1.5E-01	AW500811.1	EST_HUMAN	UI-HF-BNO-ekk-d-05-0-UI.1 NIH_MGC_50 Homo sapiens cDNA clone IMAGE:3077409 5'
8014	20952	34348	0.71	1.5E-01	U46500.1	NT	Saccharomyces cerevisiae weak multicopy suppressor of lost-1 (SOL3) gene, complete cds
8393	21362	34768	1.21	1.5E-01	P21303	SWISSPROT	MEROZONTE RECEPTOR PK68 PRECURSOR (68 KD PROTECTIVE MINOR SURFACE ANTIGEN)
8502	21530	34950	0.97	1.5E-01	AA970317.1	EST_HUMAN	cc85g12.s1 NCL_CGAP_Kd45 Homo sapiens cDNA clone IMAGE:1573030 3' similar to gb-M26062
8655	21623		1.01	1.5E-01	BE894789.1	EST_HUMAN	INTERLEUKIN-2 RECEPTOR BETA CHAIN PRECURSOR (HUMAN);
8743	21711		13.33	1.5E-01	C16800.1	EST_HUMAN	601510523F1 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:3912004 5'
8778	21743	35165	1.89	1.5E-01	L27835.1	NT	C16800 Clontech human aorta polyA+ mRNA (#6572) Homo sapiens cDNA clone GEN-5291H09 5'
8940	21808	35330	1.79	1.5E-01	D84476.1	NT	Pangasinanodon gigas growth hormone (GH) mRNA, complete cds
8982	21828		0.74	1.5E-01	P43446	SWISSPROT	Homo sapiens mRNA for ASK1, complete cds
9190	22156	35535	1.31	1.5E-01	4501972	NT	WNT-10A PROTEIN PRECURSOR
9480	22424	35882	3.06	1.5E-01	N74226.1	EST_HUMAN	Homo sapiens adaptor-related protein complex 1, beta 1 subunit (ADTB1), mRNA
9548	22511	35950	1.03	1.5E-01	BF585485.1	EST_HUMAN	za56e06.s1 Soares fetal liver spleen 1NLS Homo sapiens cDNA clone IMAGE:298868 3' similar to
9555	22517		2.73	1.5E-01	AV754819.1	EST_HUMAN	PIR-S44443 S44443 RAD23 protein homolog2 - human ;
9760	22701		0.87	1.5E-01	AI130007.1	EST_HUMAN	GYO000404 Human P-sorosis Differential Display Homo sapiens cDNA
9808	21131	34534	0.82	1.5E-01	U00455.1	NT	AV754819 TP Homo sapiens cDNA clone TPAAHB12 5'
							AI130007 NT2RP3 Homo sapiens cDNA clone NT2RP-3000080 5'
							Acipenser transmontano vitellogenin mRNA, partial cds

Page 108 of 546
Table 4

Single Exon Probes Expressed in Bone Marrow

Probe SEQ ID NO:	Exon SEQ ID NO:	ORF SEQ ID NO:	Expression Signal	Most Similar (Top) Hit BLAST E Value	Top Hit Accession No.	Top Hit Database Source	Top Hit Descriptor
10176	23101	36581	0.58	1.5E-01	M77144.1	NT	Human type II 3-beta hydroxysteroid dehydrogenase/ 5-delta - 4-delta isomerase gene, complete cds
10281	23206	36690	6.22	1.5E-01	AF007570.1	NT	Aplysia californica carboxypeptidase D mRNA, complete cds
10281	23206	36691	6.22	1.5E-01	AF007570.1	NT	Aplysia californica carboxypeptidase D mRNA, complete cds
10563	23485	36979	2.67	1.5E-01	X98852.1	NT	P. lentus scutellum mRNA for integrin beta subunit
10551	23573		0.52	1.5E-01	AB027759.1	NT	Mesocricetus auratus mRNA for collagen type XVII, complete cds
10571	23593	37089	2.49	1.5E-01	AI814046.1	EST_HUMAN	wk53h12.x1 NCI_CGAP_P22 Homo sapiens cDNA clone IMAGE:2419175 3' similar to gb:M27608 BETA
10571	23593	37089	2.49	1.5E-01	AI814046.1	EST_HUMAN	GALACTOSIDASE-RELATED PROTEIN PRECURSOR (HUMAN);
10571	23593	37089	2.49	1.5E-01	AI814046.1	EST_HUMAN	wk53h12.x1 NCI_CGAP_P22 Homo sapiens cDNA clone IMAGE:2419175 3' similar to gb:M27608 BETA
10753	23675	37172	2.19	1.5E-01	U40932.1	NT	GALACTOSIDASE-RELATED PROTEIN PRECURSOR (HUMAN);
10906	23826	37338	2.14	1.5E-01	AJ011964.1	NT	Danio rerio transcription factor Pax9b (Pax9) mRNA, complete cds
10906	23826	37338	2.14	1.5E-01	AJ011964.1	NT	Claviceps purpurea ps1 gene
10906	23826	37338	2.14	1.5E-01	AJ011964.1	NT	Claviceps purpurea ps1 gene
11173	24130	37659	4.35	1.5E-01	AL163280.2	NT	Homo sapiens chromosome 21 segment HS21C080
11173	24130	37660	4.35	1.5E-01	AL163280.2	NT	Homo sapiens chromosome 21 segment HS21C080
11412	24366	37891	1.71	1.5E-01	AW841915.1	EST_HUMAN	IL5-CN0024-030300-026-D04 CN0024 Homo sapiens cDNA
11449	24392	37897	3.86	1.5E-01	AA425488.1	EST_HUMAN	zw46d02.L1 Soares, total, testis, NB2H-IF8, 9w Homo sapiens cDNA clone IMAGE:773091 5' similar to
11501	20581	33945	1.69	1.5E-01	AB73157.1	EST_HUMAN	contains element MER22 repetitive element;
12229	25769		11.17	1.5E-01	BF700582.1	EST_HUMAN	wf62x08.x1 NCI_CGAP_UH1 Homo sapiens cDNA clone IMAGE:2491310 3'
12607	25306		1.62	1.5E-01	AF030368.2	NT	602128753F1 NIH_MGC_56 Homo sapiens cDNA clone IMAGE:4285549 5'
12610	25309		1.81	1.5E-01	AJ238332.1	NT	Rattus norvegicus chemokine CX3C mRNA, complete cds
12638	25781		12.47	1.5E-01	R83077.1	EST_HUMAN	Mus musculus mRNA for death inducer-obitator-1 (Dio-1)
12748	25814		2.45	1.5E-01	AV741272.1	EST_HUMAN	yp87x04.L1 Soares fetal liver spleen 1NFLS Homo sapiens cDNA clone IMAGE:1944390 6'
12855	25722	31613	7.87	1.5E-01	AL139074.2	NT	AV741272 CB Homo sapiens cDNA clone CBDA GD04 5'
13073	25608	31689	1.53	1.5E-01	AJ278242.1	NT	Campylobacter jejuni NCTC11168 complete genome; segment 1/8
299	13393		2.07	1.4E-01	AF009863.1	NT	Sus scrofa mRNA for sodium iodide symporter
911	13958		3.95	1.4E-01	D78638.1	NT	Homo sapiens T cell receptor beta locus, TCRBV21S2A2 region
1264	14299		2.77	1.4E-01	T91864.1	EST_HUMAN	Xenopus laevis mRNA for DNA (cytosine-5-)methyltransferase, complete cds
1763	14782		1.61	1.4E-01	6878980	NT	y454c01.s1 Soares fetal liver spleen 1NFLS Homo sapiens cDNA clone IMAGE:112032 3'
1766	14795	27780	1.82	1.4E-01	AE001710.1	NT	Mus musculus growth differentiation factor 5 (Gdf5), mRNA
2002	15023		12.82	1.4E-01	AA720615.1	EST_HUMAN	Thomomys maritima section 22 of 136 of the complete genome
2482	15488	28509	1.15	1.4E-01	P30708	SWISSPROT	my72d07.s1 NCI_CGAP_GCB1 Homo sapiens cDNA clone IMAGE:1283821 3'
2804	15798	28815	7.57	1.4E-01	AI893496.1	EST_HUMAN	GLYCEROL-3-PHOSPHATE ACYLTRANSFERASE PRECURSOR (GPAT)

Page 109 of 546
Table 4
Single Exon Probes Expressed in Bone Marrow

Probe SEQ ID NO:	Exon SEQ ID NO:	ORF SEQ ID NO:	Expression Signal	Most Similar (Top) Hit BLAST E Value	Top Hit Accession No.	Top Hit Database Source	Top Hit Descriptor
3914	18954	29885	0.93	1.4E-01	R59232.1	EST_HUMAN	y997a03.r1 Soares Infant brain 1N1B Homo sapiens cDNA clone IMAGE:41467 5'
3914	18954	29888	0.93	1.4E-01	R59232.1	EST_HUMAN	y997a03.r1 Soares Infant brain 1N1B Homo sapiens cDNA clone IMAGE:41467 5'
4202	17233	30120	11.04	1.4E-01	A1698094.1	EST_HUMAN	b56a02.x1 NCI CGAP_L124 Homo sapiens cDNA clone IMAGE:2273570 3'
4202	17233	30121	11.04	1.4E-01	A1698094.1	EST_HUMAN	b56a02.x1 NCI CGAP_L124 Homo sapiens cDNA clone IMAGE:2273570 3'
4265	17294	30173	3.96	1.4E-01	AE001710.1	NT	Thermotoga maritima section 22 of 136 of the complete genome
							2f50b01.s1 Soares fetal_liver_spleen_1N1FLS_S1 Homo sapiens cDNA clone IMAGE:463673 3' similar to gba:XD01057_maf1 INTERLEUKIN-2 RECEPTOR ALPHA CHAIN PRECURSOR (HUMAN); contains Alu repetitive element
4440	17466		0.71	1.4E-01	AA776287.1	EST_HUMAN	Homo sapiens phosphodiesterase 4A, cAMP-specific (dunce) (Drosophila)-homolog phosphodiesterase E2) (PDE4A), mRNA
4703	17724	30818	0.69	1.4E-01	5453861	NT	(PDE4A), mRNA
4907	17924	30817	0.92	1.4E-01	AV689689.1	EST_HUMAN	AV689689 GK6 Homo sapiens cDNA clone GK6DUG09 5'
5379	18483	31358	4.72	1.4E-01	T90677.1	EST_HUMAN	y915c11.s1 Stratagene lung (#637210) Homo sapiens cDNA clone IMAGE:117812 3'
5402	18506	31381	4.06	1.4E-01	AB004558.1	NT	Candida tropicalis DNA for mitochondrial NADP-linked isocitrate dehydrogenase, complete cds
5402	18506	31382	4.06	1.4E-01	AB004558.1	NT	Candida tropicalis DNA for mitochondrial NADP-linked isocitrate dehydrogenase, complete cds
6430	19498	32749	2.9	1.4E-01	BES26891.1	EST_HUMAN	h67c02.x1 NCI CGAP_K1411 Homo sapiens cDNA clone IMAGE:3133538 3'
6629	19887	32965	5.1	1.4E-01	AU117147.1	EST_HUMAN	AU117147 HEMBA1 Homo sapiens cDNA clone HEMBA1000769 5'
6829	19887	32966	5.1	1.4E-01	AU117147.1	EST_HUMAN	AU117147 HEMBA1 Homo sapiens cDNA clone HEMBA1000769 5'
6723	19779	33058	3.48	1.4E-01	AW082788.1	EST_HUMAN	h671d12.x1 Soares NFL_T_GBC_S1 Homo sapiens cDNA clone IMAGE:2581751 3'
6737	19763		1.26	1.4E-01	BE268638.1	EST_HUMAN	601169523F1 NIH_MGC_7 Homo sapiens cDNA clone IMAGE:3637581 5'
6759	19813	33093	2.02	1.4E-01	BF378533.1	EST_HUMAN	QV1-UM00336-080300-103-409 UM00336 Homo sapiens cDNA
7333	20304		0.74	1.4E-01	AL118598.1	EST_HUMAN	DKFZp761A0910_1 781 (synonym: hamy2) Homo sapiens cDNA clone DKFZp761A0910 5'
7615	20576		1.67	1.4E-01	AW015373.1	EST_HUMAN	UHH-B10-eat-c-09-Q-UI.s1 NCI CGAP_Sub1 Homo sapiens cDNA clone IMAGE:2710289 3'
7641	20601	33965	0.51	1.4E-01	F08745.1	EST_HUMAN	HSC1DB011 normalized Infant brain cDNA Homo sapiens cDNA clone c-1db01
							w04f12.x1 NCI CGAP_CL11 Homo sapiens cDNA clone IMAGE:2389295 3' similar to SW:ICE4_HUMAN
7694	20652		0.61	1.4E-01	A1762827.1	EST_HUMAN	P49882 CASPASE-4 PRECURSOR ;
7693	20627	34203	0.82	1.4E-01	U85645.1	NT	Oryctolagus cuniculus fructose 1,6, biphosphatase aldolase (AldB) gene, complete cds
8029	20968	34561	1.24	1.4E-01	A305192.1	EST_HUMAN	q30b012.x1 Soares NIH-FMPL S1 Homo sapiens cDNA clone IMAGE:1878583 3'
8818	21785		1.24	1.4E-01	AV659047.1	EST_HUMAN	AV659047 GLC Homo sapiens cDNA clone GLCFSH08 3'
							tf82b12.x1 Soares NSF_F8_9W_OT_PA_P_S1 Homo sapiens cDNA clone IMAGE:2126111 3' similar to TR:002710 O02710 GAG POLYPROTEIN ;
9136	22102		0.67	1.4E-01	A436083.1	EST_HUMAN	EST178192 Colon carcinoma (HCC) cell line Homo sapiens cDNA 5' end
9264	22230	35681	4.68	1.4E-01	AA307073.1	EST_HUMAN	qf58b03.y1 Morton Fetal Coochlea Homo sapiens cDNA clone IMAGE:2487485 5'
9345	22310	35735	0.56	1.4E-01	AW023638.1	EST_HUMAN	y10h05.r1 Soares placenta Nb2-IP Homo sapiens cDNA clone IMAGE:139873 5'
9478	22442	35883	0.94	1.4E-01	R62748.1	EST_HUMAN	y10h05.r1 Soares placenta Nb2-IP Homo sapiens cDNA clone IMAGE:139873 5'
9478	22442	35884	0.94	1.4E-01	R62748.1	EST_HUMAN	y10h05.r1 Soares placenta Nb2-IP Homo sapiens cDNA clone IMAGE:139873 5'

Page 110 of 546
Table 4

Single Exon Probes Expressed in Bone Marrow

Probe SEQ ID NO:	Exon SEQ ID NO:	ORF SEQ ID NO:	Expression Signal	Most Similar (Top) Hit BLAST E Value	Top Hit Accession No.	Top Hit Database Source	Top Hit Descriptor
8641	22504	35953	9.19	1.4E-01	BF310859.1	EST_HUMAN	601895465F1 NIH_MGC. 19 Homo sapiens cDNA clone IMAGE:4124824 5'
8630	22574	36024	1.19	1.4E-01	W09411.1	EST_HUMAN	z094e04.1 Soares fetal heart_NbH19W Homo sapiens cDNA clone IMAGE:357102 5' similar to contains element KER repetitive element ;
9702	22655	36109	0.43	1.4E-01	X73283.1	NT	M. varnelli genes rpoH, rpoB and rpoA
9702	22655	36110	0.43	1.4E-01	X73283.1	NT	M. varnelli genes rpoH, rpoB and rpoA
9713	22668	36123	1.44	1.4E-01	Y10198.1	NT	Homo sapiens PH6X gene
9713	22668	36124	1.44	1.4E-01	Y10198.1	NT	Homo sapiens PH6X gene
9805	21128	34532	2.08	1.4E-01	AF121361.1	NT	Drosophila melanogaster signal transducing adaptor protein (STAM), serine threonine kinase Ial (IAL), and zinc finger protein (DNZ1) genes, complete cds
10164	23089	36557	0.55	1.4E-01	X68092.1	NT	C. perfingens ORF for putative membrane transport protein
10346	23270	36747	1.12	1.4E-01	AF023813.1	NT	Macromitrium levalium small ribosomal protein 4 (psr4) gene, chloroplast gene encoding chloroplast protein, partial cds
10448	23370	36861	0.57	1.4E-01	AW021908.1	EST_HUMAN	af23h08.y1 Morton Fetal Cochlea Homo sapiens cDNA clone IMAGE:2485094 5'
10448	23370	36862	0.57	1.4E-01	AW021908.1	EST_HUMAN	af23h08.y1 Morton Fetal Cochlea Homo sapiens cDNA clone IMAGE:2485094 5'
10619	23541	37039	0.87	1.4E-01	BF376285.1	EST_HUMAN	MR3-ST0218-211289-013-a08 ST0218 Homo sapiens cDNA
10619	23541	37040	0.87	1.4E-01	BF376285.1	EST_HUMAN	MR3-ST0218-211289-013-a08 ST0218 Homo sapiens cDNA
10829	23750		0.56	1.4E-01	T84283.1	EST_HUMAN	y447d03.1 Soares fetal liver spleen 1NFLS Homo sapiens cDNA clone IMAGE:111365 5'
10952	23882	37395	0.71	1.4E-01	Z89117.1	NT	Bacillus subtilis complete genome (section 14 of 21); from 2589451 to 2812870
11191	24147	37680	2.59	1.4E-01	R63400.1	EST_HUMAN	y70c05.1 Soares breast 2NtH18t Homo sapiens cDNA clone IMAGE:154088 5'
11432	24376	37918	2.53	1.4E-01	P08848	SWISSPROT	INTEGRIN ALPHA-5 PRECURSOR (FIBRONECTIN RECEPTOR ALPHA SUBUNIT) (INTEGRIN ALPHA-F) (VLA-5) (CD49E)
11762	24637		1.59	1.4E-01	AL161408.2	NT	Arabidopsis thaliana DNA chromosome 4, contig fragment No. 8
11797	23952	37474	2.38	1.4E-01	U28780.1	NT	Borrelia burgdorferi glyceraldehyde-3-phosphate dehydrogenase (GAPDH), phosphoglycerate kinase (PGK),
11855	24737		1.55	1.4E-01	X82102.1	NT	triosephosphate isomerase (TPI) genes, complete cds
12549	25272	31778	2.33	1.4E-01	X74773.1	NT	M. musculus p16K gene for 16 kDa protein
12562	25280		2.24	1.4E-01		NT	P. salina plastid gene secY
12605	25984		2.35	1.4E-01	BE513802.1	EST_HUMAN	Rattus norvegicus desmin (Des), mRNA
12702	26382		2.29	1.4E-01	D84004.1	NT	601315638F1 NIH_MGC. 8 Homo sapiens cDNA clone IMAGE:3684329 5'
12776	25981		4.88	1.4E-01	P10447	SWISSPROT	Synechocystis sp. PCC6803 complete genome, 23/27, 2888767-3002965
12998	25792		3.72	1.4E-01	D82983.1	NT	TYROSINE-PROTEIN KINASE TRANSFORMING PROTEIN ABL
13087	25601		2.63	1.4E-01	AW377698.1	EST_HUMAN	Mus musculus mRNA for prolifase, complete cds
322	13414	26338	2.48	1.3E-01	4758467	NT	MR0-HT0208-221289-204-c08 HT0208 Homo sapiens cDNA
322	13414	26339	2.48	1.3E-01	4758467	NT	Homo sapiens G protein-coupled receptor 50 (GPR50) mRNA

Table 4

Single Exon Probes Expressed in Bone Marrow.

Probe SEQ ID NO:	Exon SEQ ID NO:	ORF SEQ ID NO:	Expression Signal	Most Similar (Top) Hit BLAST E Value	Top Hit Accession No.	Top Hit Database Source	Top Hit Descriptor
530	13601	26619	2.91	1.3E-01	AB013138.1	NT	Homo sapiens gene for NBS1, complete cds
637	13703	26624	0.78	1.3E-01	AJ277606.1	NT	Human calicivirus HUNLV/Girlington/83/UK RNA for capsid protein (ORF2), strain HUNLV/Girlington/83/UK
637	13703	26625	0.78	1.3E-01	AJ277606.1	NT	Human calicivirus HUNLV/Girlington/83/UK RNA for capsid protein (ORF2), strain HUNLV/Girlington/83/UK
845	13901	26659	1.04	1.3E-01	X63330.1	NT	P. dumerilii histone gene cluster for core histones H2A, H2B, H3 and H4
895	13980	26808	1.76	1.3E-01	AF139518.1	NT	Rattus norvegicus A-kinase anchor protein mRNA, complete cds
1028	14074	27024	1.8	1.3E-01	AL117078.1	NT	Botrytis cinerea strain T4 cDNA library under conditions of nitrogen deprivation
1129	14172		2.88	1.3E-01	AL116285.1	NT	Botrytis cinerea strain T4 cDNA library under conditions of nitrogen deprivation
1220	14288	27215	1.51	1.3E-01	AV712467.1	EST_HUMAN	AV712467 DCA Homo sapiens cDNA clone DCAAFF05 5'
1438	14471		0.93	1.3E-01	AF146277.1	NT	Homo sapiens adapter protein GMS mRNA, complete cds
1978	14987	27999	2.07	1.3E-01	AL117078.1	NT	Botrytis cinerea strain T4 cDNA library under conditions of nitrogen deprivation
2180	15195		1.69	1.3E-01	AJ243578.1	NT	Rhodospirillum rubrum acidophilus pucB5, pucA6, pucB6, pucA7, pucB8, pucA8 and pucC genes and ORF151
2298	15310		0.96	1.3E-01	AW612104.1	EST_HUMAN	RC4-ST0173-101099-032-d12 ST0173 Homo sapiens cDNA
2389	15397		3.94	1.3E-01	AE001016.1	NT	Archaeoglobus fulgidus section 91 of 172 of the complete genome
2592	15593	28810	2.23	1.3E-01	M89918.1	NT	Caraassius auratus keratin type I mRNA, complete cds
3465	16511	29432	0.98	1.3E-01	M1672.1	NT	Bovine branched chain alpha-keto acid dihydrolipoyl transacylase mRNA, complete cds
3735	16777	29689	0.91	1.3E-01	AP000001.1	NT	Pyrococcus horikoshii OT3 genomic DNA, 1-287000 nt. position (1/7)
3735	16777	29690	0.91	1.3E-01	AP000001.1	NT	Pyrococcus horikoshii OT3 genomic DNA, 1-287000 nt. position (1/7)
3741	16783	29695	0.89	1.3E-01	AB032188.1	NT	Homo sapiens DD4 gene for dihydrolipoyl dehydrogenase 4 [AKR 1C4], exon 2
3794	16777	29689	0.67	1.3E-01	AP000001.1	NT	Pyrococcus horikoshii OT3 genomic DNA, 1-287000 nt. position (1/7)
3794	16777	29690	0.67	1.3E-01	AP000001.1	NT	Pyrococcus horikoshii OT3 genomic DNA, 1-287000 nt. position (1/7)
3816	16958	29762	0.74	1.3E-01	6978940	NT	Rattus norvegicus Fibrinogen, gamma polypeptide (F99), mRNA
4014	17053		1.48	1.3E-01	AL161581.2	NT	Arabidopsis thaliana DNA chromosome 4, contig fragment No. 77
4081	13703	26624	7.11	1.3E-01	AJ277606.1	NT	Human calicivirus HUNLV/Girlington/83/UK RNA for capsid protein (ORF2), strain HUNLV/Girlington/83/UK
4081	13703	26625	7.11	1.3E-01	AJ277606.1	NT	Human calicivirus HUNLV/Girlington/83/UK RNA for capsid protein (ORF2), strain HUNLV/Girlington/83/UK
4163	17194		1.12	1.3E-01	AF020713.1	NT	Bacteriophage SPBc2 complete genome
4185	17216		4.1	1.3E-01	AW364341.1	EST_HUMAN	QV3-DT0018-081299-036-ab3 DT0018 Homo sapiens cDNA
4194	17225	30114	2.47	1.3E-01	AF026805.1	NT	Schistosoma mansoni fructose biphosphate aldolase mRNA, complete cds
4215	17244	30128	20.19	1.3E-01	AW273741.1	EST_HUMAN	x22310.1 Soares_NFL_T_GBC_S1 Homo sapiens cDNA clone IMAGE-2813986 3'
4343	17370		1.49	1.3E-01	AL163280.2	NT	Homo sapiens chromosome 21 segment HS21C080

Page 112 of 546
Table 4

Single Exon Probes Expressed in Bone Marrow

Probe SEQ ID NO:	Exon SEQ ID NO:	ORF SEQ ID NO:	Expression Signal	Most Similar (Top) Hit BLAST E Value	Top Hit Accession No.	Top Hit Database Source	Top Hit Descriptor
4505	17530	30414	0.8	1.3E-01	M21572.1	NT	Bovine branched chain alpha-keto acid dihydrolylpyl transacylase mRNA, complete cds
4564	17587	30479	2.72	1.3E-01	BE272339.1	EST_HUMAN	601126096F1 NIH_MGC_9 Homo sapiens cDNA clone IMAGE:2890063 5'
5014	18028		0.91	1.3E-01	AU136619.1	EST_HUMAN	AU136619 PLACE1 Homo sapiens cDNA clone PLACE1004663 5'
5080	18070		0.63	1.3E-01	BF091980.1	EST_HUMAN	RC4-TN0077-180900-012-c05 TN0077 Homo sapiens cDNA
5388	18501	31379	0.78	1.3E-01	AW468888.1	EST_HUMAN	ha07006.x1 NCI_CGAP_Kid12 Homo sapiens cDNA clone IMAGE:2872879 3' similar to contains L1.b1 L1
5438	18538	31448	1.98	1.3E-01	AW804417.1	EST_HUMAN	L1 repetitive element
5579	18675		0.91	1.3E-01	AF107783.1	NT	QVQ-UM0083-100400-189-a06 UM0083 Homo sapiens cDNA
5665	18760		0.68	1.3E-01	AF056880.1	NT	Ern1c2a nidulans DNA-dependent RNA polymerase II RPB140 (RPB2) gene, partial cds
5813	18803	32086	0.85	1.3E-01	BF210920.1	EST_HUMAN	Hepatitis C virus 88_CL10 genome polyprotein gene, partial cds
6088	19178	32386	0.58	1.3E-01	BF527281.1	EST_HUMAN	601874591F1 NIH_MGC_54 Homo sapiens cDNA clone IMAGE:4101119 5'
6089	19178	32397	0.58	1.3E-01	BF527281.1	EST_HUMAN	602039337F2 NCI_CGAP_Bim67 Homo sapiens cDNA clone IMAGE:4177233 5'
6330	19688	32987	17.29	1.3E-01	AB031328.1	NT	602039337F2 NCI_CGAP_Bim67 Homo sapiens cDNA clone IMAGE:4177233 5'
6720	19776	33055	2.08	1.3E-01	X88891.1	NT	Schizosaccharomyces pombe gene for Alp41, complete cds
6940	20184	33487	0.64	1.3E-01	HF73425.1	EST_HUMAN	C.jecchus Intron 4 of visual pigment gene (red allele)
6955	20180		0.82	1.3E-01	W28367.1	EST_HUMAN	yu02401.1 Soares fetal liver spleen 1NFLS Homo sapiens cDNA clone IMAGE:232609 5'
7008	20135	33450	1.04	1.3E-01	BE782928.1	EST_HUMAN	2683 Human retina cDNA randomly primed sublibrary Homo sapiens cDNA
7009	20135	33451	1.04	1.3E-01	BE782928.1	EST_HUMAN	601465957F1 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:3869079 5'
7211	20234		0.72	1.3E-01	BF529560.1	EST_HUMAN	601465957F1 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:3869079 5'
7477	20443		2.15	1.3E-01	H48684.1	EST_HUMAN	602044345F1 NCI_CGAP_Bim67 Homo sapiens cDNA clone IMAGE:4181868 5'
8292	21261		0.97	1.3E-01	BE272339.1	EST_HUMAN	6036021.1 Soares fetal liver spleen 1NFLS Homo sapiens cDNA clone IMAGE:207076 5'
8308	21276	34688	1.59	1.3E-01	11423294	NT	601128096F1 NIH_MGC_9 Homo sapiens cDNA clone IMAGE:2890063 5'
8336	21305	34720	1.18	1.3E-01	BF690622.1	EST_HUMAN	Homo sapiens PRO0811 protein (PRO0811), mRNA
8580	21548		0.56	1.3E-01	BE682528.1	EST_HUMAN	602187015T1 NIH_MGC_49 Homo sapiens cDNA clone IMAGE:4289074 3'
8617	21685	35001	0.8	1.3E-01	11421558	NT	60133828F1 NIH_MGC_44 Homo sapiens cDNA clone IMAGE:3689834 5'
8688	21656		4.68	1.3E-01	Z74102.1	NT	Homo sapiens TED protein (TED), mRNA
8729	21697		4.2	1.3E-01	8823918	NT	S. cerevisiae chromosome IV reading frame ORF YDL054c
8873	21840	35283	1.05	1.3E-01	BF680522.1	EST_HUMAN	Homo sapiens core histone macroH2A2.2 (MACROH2A2), mRNA
8900	22265	35684	0.55	1.3E-01	R11172.1	EST_HUMAN	602187015T1 NIH_MGC_49 Homo sapiens cDNA clone IMAGE:4289074 3'
9300	22285	35695	0.55	1.3E-01	R11172.1	EST_HUMAN	y39g11.1.1 Soares fetal liver spleen 1NFLS Homo sapiens cDNA clone IMAGE:289076 5'
9574	22536	35987	0.65	1.3E-01	11068003	NT	SP:RL2B RAT P29316 60S RIBOSOMAL PROTEIN ;
9674	22536	35988	0.65	1.3E-01	11068003	NT	y39g11.1.1 Soares fetal liver spleen 1NFLS Homo sapiens cDNA clone IMAGE:129284 5' similar to
							SP:RL2B RAT P29316 60S RIBOSOMAL PROTEIN ;
							y39g11.1.1 Soares fetal liver spleen 1NFLS Homo sapiens cDNA clone IMAGE:129284 5' similar to
							Plutella xylostella granulovirus, complete genome
							Plutella xylostella granulovirus, complete genome

Table 4

Single Exon Probes Expressed in Bone Marrow

Probe SEQ ID NO:	Exon SEQ ID NO:	ORF SEQ ID NO:	Expression Signal	Most Similar (Top) Hit BLAST E Value	Top Hit Accession No.	Top Hit Database Source	Top Hit Descriptor
9827	22676	36132	3.9	1.3E-01	AF023128.1	NT	Oryzotegus curticulus H+K+ATPase alpha 2b subunit mRNA, complete cds
10128	23055		0.74	1.3E-01	N86348.1	EST_HUMAN	J7837F Human fetal heart, Lambda ZAP Express Homo sapiens cDNA clone J7837 5' similar to B-CELL
10411	23333		0.89	1.3E-01	8393940	NT	RECEPTOR ASSOCIATED PROTEIN (BAP) 29
10489	23411	36908	0.86	1.3E-01	AW951588.1	EST_HUMAN	Rattus norvegicus peptidyl arginine deiminase, type IV (Pdli4), mRNA
10757	25702	37173	1.06	1.3E-01	AL163246.2	NT	MF2-CT0222-201089-001-e01 CT0222 Homo sapiens cDNA
10891	23811	37318	0.86	1.3E-01	AU121237.1	EST_HUMAN	Homo sapiens chromosome 21 segment HS21C048
10898	23856	37372	0.45	1.3E-01	AW247836.1	EST_HUMAN	AU121237 HEMBB1 Homo sapiens cDNA clone HEMBB1002387 5'
10896	23862		2.57	1.3E-01	BF330899.1	EST_HUMAN	2820637.3prime NIH_MGC 7 Homo sapiens cDNA clone IMAGE:2820637 3'
11581	24519		3.56	1.3E-01	6671745	NT	MR4-BT0358-130700-010-H08 BT0358 Homo sapiens cDNA
11670	24608	38182	1.77	1.3E-01	AW082836.1	EST_HUMAN	Mus musculus cofilin 2, muscle (Cf2), mRNA
11670	24608	38183	1.77	1.3E-01	AW082836.1	EST_HUMAN	xc20709.x1 NCL_CGAP_Cot19 Homo sapiens cDNA clone IMAGE:2584841 3'
11622	24803	38395	2.33	1.3E-01	BE279448.1	EST_HUMAN	xc20709.x1 NCL_CGAP_Cot19 Homo sapiens cDNA clone IMAGE:2584841 3'
12397	25171	31817	1.83	1.3E-01	BE618346.1	EST_HUMAN	601158052F1 NIH_MGC 21 Homo sapiens cDNA clone IMAGE:3504804 5'
12335	25260		3.3	1.3E-01	AL242790.1	NT	601462741F1 NIH_MGC 87 Homo sapiens cDNA clone IMAGE:3968003 5'
12692	25357	31766	1.37	1.3E-01	BF572303.1	EST_HUMAN	Gallus gallus sox1 gene for lymphotactin, exons 1-3
12885	25475		1.48	1.3E-01	AB026829.1	NT	Ephradytia fundalis mRNA for sALK-6, complete cds
12915	25494		2.52	1.3E-01	AW001114.1	EST_HUMAN	wu24409.x1 Soares_Dieckgrasie, colon_NHCD Homo sapiens cDNA clone IMAGE:2520877 3' similar to TR:O80287 O60287 KIAA0539 PROTEIN. ;
383	13488	26428	14.86	1.2E-01	AK421744.1	EST_HUMAN	U99002.x1 NCL_CGAP_Bm23 Homo sapiens cDNA clone IMAGE:2098539 3' similar to gb:U05760_mai
424	13119		1.38	1.2E-01	U66912.1	NT	ANNEXIN V (HUMAN);
549	13819		2.78	1.2E-01	AF039442.1	NT	Dicystosellum discoideum ORF DG1016 gene, partial cds
1377	14411	27381	2.19	1.2E-01	AU149146.1	EST_HUMAN	Homo sapiens colon cancer antigen NY-CO-48 mRNA, partial cds
1377	14411	27382	2.19	1.2E-01	AU149146.1	EST_HUMAN	AU149146 NT2RM4 Homo sapiens cDNA clone NT2RM4001691 3'
1383	14417		4.89	1.2E-01	AV735249.1	EST_HUMAN	AU149148 NT2RM4 Homo sapiens cDNA clone NT2RM4001691 3'
1388	14421		0.92	1.2E-01	AL445098.1	NT	AV735249 cda Homo sapiens cDNA clone cdaAJB11 5'
1607	14540		1.23	1.2E-01	AA897474.1	EST_HUMAN	Thermoplasma acidophilum complete genome, segment 4/5
1635	14657	27643	1.62	1.2E-01	Q14694	SWISSPROT	a448609.s1 Soares_NFL_T_GBC_S1 Homo sapiens cDNA clone IMAGE:1460584 3' similar to TR:Q16671
1638	14690	27698	2.89	1.2E-01	AL285402.1	EST_HUMAN	Q16671 ANTI-MULLERIAN HORMONE TYPE II RECEPTOR PRECURSOR. ;
1783	14812		18.69	1.2E-01	X89211.1	NT	NUCLEAR FACTOR OF ACTIVATED T-CELLS, CYTOPLASMIC 4 (T CELL TRANSCRIPTION FACTOR
1634	14953		1.76	1.2E-01	AW449368.1	EST_HUMAN	MFAT93(NF-ATC4) (NF-AT3)
							q168009.x1 NCL_CGAP_Eco2 Homo sapiens cDNA clone IMAGE:1960553 3'
							H.sapiens DNA for endogenous retroviral like element
							UIH-BI3-ald-e-10-Q-UJ.s1 NCL_CGAP_Sub55 Homo sapiens cDNA clone IMAGE:2734654 3'

Table 4

Single Exon Probes Expressed in Bone Marrow

Probe SEQ ID NO:	Exon SEQ ID NO:	ORF SEQ ID NO:	Expression Signal	Most Similar (Top) Hit BLAST E Value	Top Hit Accession No.	Top Hit Database Source	Top Hit Descriptor
2193	15208	28227	1.65	1.2E-01	BF248490.1	EST_HUMAN	601821567F1 NIH_MGC_62 Homo sapiens cDNA clone IMAGE:4048224 5'
2294	15308	28329	1.21	1.2E-01	AL163213.2	NT	Homo sapiens chromosome 21 segment HS21C013
2595	15598	28814	2.05	1.2E-01	AW086568.1	EST_HUMAN	QV3-BN0046-220300-129-F10 BN0046 Homo sapiens cDNA
2602	15602	28824	1.61	1.2E-01	BE218889.1	EST_HUMAN	hV65704.x1 NCI_CGAP_Lu24 Homo sapiens cDNA clone IMAGE:3178303 3'
							ts18g07.x1 NCI_CGAP_Pan1 Homo sapiens cDNA clone IMAGE:2228888 3' similar to TRQ14048 Q14048 COLLAGEN VI ALPHA-2 ALTERNATIVE C-TERMINAL DOMAIN. [1] ;contains element PTR5 repetitive element;
2738	15732	28748	37.59	1.2E-01	AI623388.1	EST_HUMAN	Human E1A enhancer binding protein (E1A-F) mRNA, partial cds
2855	15915	28838	1.4	1.2E-01	U18018.1	NT	aa80c09.x1 Barstead color HPLRB7 Homo sapiens cDNA clone IMAGE:2335024 3' similar to gbl.05085
2813	15971	28885	2.03	1.2E-01	AI720470.1	EST_HUMAN	60S RIBOSOMAL PROTEIN L30 (HUMAN); Human creatine kinase-B mRNA, complete cds
2946	16004	28829	2.52	1.2E-01	M16364.1	NT	Wheat mRNA for a group 3 late embryogenesis abundant protein (LEA)
3017	16075	28998	0.73	1.2E-01	X56882.1	NT	QV1-BT0259-261039-021-055 BT0259 Homo sapiens cDNA
3244	16289	29224	1.34	1.2E-01	AW370688.1	EST_HUMAN	Methanococcus jannaschii section 142 of 150 of the complete genome
3271	16325		0.82	1.2E-01	U67600.1	NT	Wheat mRNA for a group 3 late embryogenesis abundant protein (LEA)
3533	16579	29502	0.86	1.2E-01	X56882.1	NT	Wheat mRNA for a group 3 late embryogenesis abundant protein (LEA)
3533	16579	29503	0.86	1.2E-01	X56882.1	NT	Bacillus subtilis complete genome (section 15 of 21); from 2795131 to 3013540
3620	16538		1.05	1.2E-01	Z99118.1	NT	601810786R1 NIH_MGC_46 Homo sapiens cDNA clone IMAGE:4053688 3'
3780	16821		0.7	1.2E-01	BF128551.1	EST_HUMAN	P.clarkii mRNA; repeat region (ID 2MR17)
4211	17240	30125	2.16	1.2E-01	Z64256.1	NT	P.clarkii mRNA; repeat region (ID 2MR17)
4211	17240	30126	2.16	1.2E-01	Z64256.1	NT	Lescentium mRNA for glyceralde-1
4751	17771	30867	0.98	1.2E-01	Z48183.1	NT	HEMOLYSIN PRECURSOR
5116	18128		1	1.2E-01	P16468	SWISSPROT	Homo sapiens chromosome 21 segment HS21C027
5150	18159	31038	0.91	1.2E-01	AL163227.2	NT	Homo sapiens chromosome 21 segment HS21C027
5150	18159	31039	0.91	1.2E-01	AL163227.2	NT	Arabidopsis thaliana DNA chromosome 4, contig fragment No. 30
5164	18173	31052	1	1.2E-01	AL161618.2	NT	601680493R2 NIH_MGC_83 Homo sapiens cDNA clone IMAGE:3950711 3'
5234	18242	31114	0.81	1.2E-01	BE974502.1	EST_HUMAN	m63c04.s1 NCI_CGAP_GCB1 Homo sapiens cDNA clone IMAGE:1282850 3'
5322	18428	31178	0.75	1.2E-01	AA744368.1	EST_HUMAN	Homo sapiens calcium channel alpha1E subunit (CACNA1E) gene, exons 7-49, and partial cds, alternatively spliced
5373	18478	31351	0.91	1.2E-01	AF223391.1	NT	zcd08cd02.r1 Soares_parathyroid_tumor_Nb1IPA Homo sapiens cDNA clone IMAGE:321698 5'
5383	18497	31361	2.27	1.2E-01	W33035.1	EST_HUMAN	Homo sapiens gene encoding plakophilin (exons 1-13)
5442	18544	31458	2.15	1.2E-01	Z98283.1	NT	M.domestica Borkh. Granny Smith adh mRNA for alcohol dehydrogenase
5583	18679	31643	0.95	1.2E-01	Z48234.1	NT	601493518F1 NIH_MGC_70 Homo sapiens cDNA clone IMAGE:3885613 5'
6324	19394	32536	2.66	1.2E-01	BE620945.1	EST_HUMAN	MATING-TYPE P-SPECIFIC POLYPEPTIDE PI
6376	19444	32688	1.1	1.2E-01	P10842	SWISSPROT	

Table 4

Single Exon Probes Expressed in Bone Marrow

Probe SEQ ID NO:	Exon SEQ ID NO:	ORF SEQ ID NO:	Expression Signal	Most Similar (Top) Hit BLAST E Value	Top Hit Accession No.	Top Hit Database Source	Top Hit Descriptor
6431	19497	32750	2.19	1.2E-01	AW845275.1	EST_HUMAN	IL0-CT0031-221089-113-e04 CT0031 Homo sapiens cDNA
6498	19553	32815	1.38	1.2E-01	M268926.1	NT	Mouse galactosyltransferase mRNA, complete cds
6804	19858	33146	0.89	1.2E-01	BF347885.1	EST_HUMAN	602023112F1 NCI_CGAP_Bn07 Homo sapiens cDNA clone IMAGE:4158388 5'
6968	20180	33517	0.55	1.2E-01	AF295739.1	NT	JC virus agnoprotein, VP2, VP3, VP1, large T antigen, and small t antigen genes, complete cds
7210	20233	33587	0.87	1.2E-01	H47769.1	EST_HUMAN	yp80f04.J1 Soares fetal liver spleen 1NFLS Homo sapiens cDNA clone IMAGE:183758 5'
7210	20233	33588	0.87	1.2E-01	H47769.1	EST_HUMAN	yp80f04.J1 Soares fetal liver spleen 1NFLS Homo sapiens cDNA clone IMAGE:183758 5'
7854	20789	34175	0.82	1.2E-01	AJ271741.1	NT	Homo sapiens partial ILF3 gene for interleukin enhancer binding factor 3 (alternative transcripts dbp76, dbp78 gamma, dbp78 alpha and ILF3)
8063	21000	34398	0.9	1.2E-01	BF880613.1	EST_HUMAN	602156195F1 NIH_MGC 83 Homo sapiens cDNA clone IMAGE:4286382 5'
8088	21034	34433	0.57	1.2E-01	D87488.1	NT	Human mRNA for KIAA0282 gene, partial cds
8088	21034	34434	0.57	1.2E-01	D87488.1	NT	Human mRNA for KIAA0282 gene, partial cds
8224	21183		1.24	1.2E-01	BE007072.1	EST_HUMAN	PM3-BN0137-250300-002-f09 BN0137 Homo sapiens cDNA
8285	21204	34875	3.1	1.2E-01	AI813753.1	EST_HUMAN	wc58g03.x1 NCI_CGAP_C03 Homo sapiens cDNA clone IMAGE:2326804 3' similar to SW-GST2_HUMAN
8341	21310	34724	0.76	1.2E-01	Q02369	SWISSPROT	Q88735 MICROSMAL GLUTATHIONE S-TRANSFERASE II;
8652	21620	35040	0.81	1.2E-01	AI832881.1	EST_HUMAN	sl71b10.x1 Barstead colon HPLRB7 Homo sapiens cDNA clone IMAGE:2377435 3'
8739	21707		10.85	1.2E-01	AW083652.1	EST_HUMAN	xc49d07.x1 NCI_CGAP_Esc2 Homo sapiens cDNA clone IMAGE:2687597 3' similar to gb:M13452 LAMIN A (HUMAN);
8759	21726						Staphylococcus aureus plasmid pSK23 putative recombinase Sin (sin) gene, partial cds; and transcriptional regulator QacR (qacR) and multidrug efflux protein QacB (qacB) genes, complete cds
8798	21765	36187	3.98	1.2E-01	AF053772.1	NT	N.crassa vacuolar ATPase 57-Kd subunit (vma-2) gene, complete cds
8798	21765	35188	0.89	1.2E-01	J03858.1	NT	N.crassa vacuolar ATPase 57-Kd subunit (vma-2) gene, complete cds
8848	21814		0.81	1.2E-01	AJ271738.1	NT	Homo sapiens Xq pseudocentromeric region, segment 2/2
9037	22003		2.01	1.2E-01	U32714.1	NT	Haemophilus influenzae Rd section 29 of 163 of the complete genome
9073	22039		0.84	1.2E-01	X15191.1	NT	M.musculus DNA fragment of Apolipoprotein B gene
9830	22813	36287	1.69	1.2E-01	X77981.1	NT	S.cerevisiae HXT5 gene
10365	23288	36785	1.64	1.2E-01	AV710857.1	EST_HUMAN	AV710857 Cu Homo sapiens cDNA clone CUAKE08 5'
11232	24185		2.36	1.2E-01	D26184.1	NT	Yeast MPT5 gene for suppressor protein, complete cds
11404	24348		2.55	1.2E-01	BE862324.2	EST_HUMAN	601655578R1 NIH_MGC 65 Homo sapiens cDNA clone IMAGE:3846283 3'
11482	24425		1.57	1.2E-01	BF314481.1	EST_HUMAN	601900763F1 NIH_MGC 19 Homo sapiens cDNA clone IMAGE:4130103 5'
11584	24522	38077	2.45	1.2E-01	AF190483.1	NT	Homo sapiens dyx19 intermediate chain DNAI1 (DNAI1) gene, exon 17
11738	24823	38201	1.8	1.2E-01	9894174	NT	Homo sapiens UDP-Gal4:beta-GalNAc beta 1,4-galactosyltransferase, polypeptide 4 (B4GALT4), mRNA

Table 4

Single Exon Probes Expressed in Bone Marrow

Probe SEQ ID NO:	Exon SEQ ID NO:	ORF SEQ ID NO:	Expression Signal	Most Similar (Top) Hit BLAST E Value	Top Hit Accession No.	Top Hit Database Source	Top Hit Description
11836	24719		1.54	1.2E-01	M85109.1	NT	Rabbit glycogen-associated protein phosphatase regulatory subunit (RG1) mRNA, complete cds
12162	25016		3.68	1.2E-01	AV658033.1	EST_HUMAN	AV658033 GLG Homo sapiens cDNA clone GLCFB12 3'
12517	25248		2.78	1.2E-01	AJ271738.1	NT	Homo sapiens Xq pseudautosomal region; segment 272
12583	25922	31304	6.17	1.2E-01	Q04912	SWISSPROT	MACROPHAGE-STIMULATING PROTEIN RECEPTOR PRECURSOR (MSP RECEPTOR) (P185-RON) (CDW136) (CD136 ANTIGEN)
12805	25432		2.68	1.2E-01	X53981.1	NT	R. norvegicus NF88 gene for 68kDa neurofilament
12867	25935	31310	1.58	1.2E-01	BE061418.1	EST_HUMAN	QV4-BT0234-111199-031-g10 BT0234 Homo sapiens cDNA
12889	25477	31731	9.62	1.2E-01	A1288803.1	EST_HUMAN	qn20g05.x1 NCI CGAP_Lu5 Homo sapiens cDNA clone IMAGE:1888840 3'
12911	25490		2.91	1.2E-01	L10187.1	NT	Xenopus laevis Integrin alpha 3 subunit mRNA, partial cds
12917	25882		9.28	1.2E-01	O88433	SWISSPROT	CYGLIN T
12946	25525	31712	1.78	1.2E-01	AE004428.1	NT	Vibrio cholerae chromosome II, section 85 of the complete chromosome
13092	16538		2.08	1.2E-01	Z89118.1	NT	Bacillus subtilis complete genome (section 15 of 21): from 2795131 to 3013540
13098	25623		1.5	1.2E-01	BF314491.1	EST_HUMAN	601900763F1 NIH_MGC_19 Homo sapiens cDNA clone IMAGE:4130103 5'
586	13636	26552	1.01	1.1E-01	A1581003.1	EST_HUMAN	in18d08.x1 NCI CGAP_Bm25 Homo sapiens cDNA clone IMAGE:2167983 3'
617	13682	26589	1.84	1.1E-01	AA569006.1	EST_HUMAN	nm008g11.s1 NCI CGAP_Co10 Homo sapiens cDNA clone IMAGE:1059620 3' similar to gb:X08895_ma1
1057	14103	27054	1.56	1.1E-01	BF687308.1	EST_HUMAN	HEME OXYGENASE 1 (HUMAN);
1087	14131		1.48	1.1E-01	AL161580.2	NT	602129847F1 NIH_MGC_58 Homo sapiens cDNA clone IMAGE:4288771 5'
1163	15930	27158	4.6	1.1E-01	AW972158.1	EST_HUMAN	Arabidopsis thaliana DNA chromosome 4, contig fragment No. 80
1254	14280	27255	2.03	1.1E-01	D64004.1	NT	EST384142 IMAGE resequences, MAGL Homo sapiens cDNA
1524	14558	27527	2.48	1.1E-01	AU140363.1	EST_HUMAN	Synechocystis sp. PCC6803 complete genome, 23/27, 2888767-3002865
2322	15333		2.72	1.1E-01	6758215	NT	AU140363 PLACE2 Homo sapiens cDNA clone PLACE2000403 5'
2545	15828		0.93	1.1E-01	6978876	NT	Mus musculus pre T-cell antigen receptor alpha (Ptra), mRNA
2574	15575		0.93	1.1E-01	AW821909.1	EST_HUMAN	Rattus norvegicus Procollagen II alpha 1 (Col2a1), mRNA
3046	16103	28017	0.95	1.1E-01	F03265.1	EST_HUMAN	RCO-ST0378-210100-032-g04 ST0379 Homo sapiens cDNA
3362	16403		1.68	1.1E-01	6763231	NT	HSC1RF022 normalized infant brain cDNA Homo sapiens cDNA clone c-1rf02 3'
3432	16480	28359	2.18	1.1E-01	BE393193.1	EST_HUMAN	Mus musculus calcium channel, voltage-dependent, T type, alpha 1G subunit (Cacna1g), mRNA
3466	16512	29433	1.3	1.1E-01	X62135.1	NT	601308978F1 NIH_MGC_44 Homo sapiens cDNA clone IMAGE:3627066 5'
3596	16841	28560	0.8	1.1E-01	Y07695.1	NT	C.reinhardtii nuclear gene on linkage group XIX
3715	16758		0.86	1.1E-01	P97384	SWISSPROT	A. Imnerius gene for transposase
3722	16766	29676	1.23	1.1E-01	X52708.1	NT	ANNEXIN XI (CALCYCLIN-ASSOCIATED ANNEXIN 50) (CAP-50)
4137	17169	30054	1.14	1.1E-01	AW819412.1	EST_HUMAN	G.gallus gene encoding non-histone chromosomal protein HM(G-14b, exons 4 and 5
4137	17169	30055	1.14	1.1E-01	AW819412.1	EST_HUMAN	MF3-ST0290-290100-025-g07 ST0290 Homo sapiens cDNA
4281	17310		9.36	1.1E-01	AF157068.1	NT	MF3-ST0290-290100-025-g07 ST0290 Homo sapiens cDNA
							Drosophila melanogaster Karscht protein (Klar) mRNA, complete cds

Page 117 of 546
Table 4
Single Exon Probes Expressed in Bone Marrow

Probe SEQ ID NO:	Exon SEQ ID NO:	ORF SEQ ID NO:	Expression Signal	Most Similar (Top) Hit BLAST E Value	Top Hit Accession No.	Top Hit Database Source	Top Hit Descriptor
4316	17345	30228	0.69	1.1E-01	AW802058.1	EST_HUMAN	IL5-UM0070-020500-069-c08 UM0070 Homo sapiens cDNA
4670	17891	30577	1.02	1.1E-01	S44867.1	NT	Tapa-1=Integral membrane protein TAPA-1 [mouse, B cell lymphoma line 38C13, Genomic, 1973 nt, segment 1 of 7]
4868	17888	30774	0.98	1.1E-01	Y07885.1	NT	A. litmarinus gene for transposase
5056	17174		0.76	1.1E-01	AF030001.1	NT	Mus musculus major histocompatibility locus class III region/butyrophilin-like protein gene, partial cds; Notch4, PBX2, RAGE, lysophosphatidic acid acyl transferase-alpha, palmitoyl-protein thioesterase 2 (PPT2), CREB-RP, and tenascin X (TNX) genes, complete
5763	18847		1.36	1.1E-01	AA747218.1	EST_HUMAN	nc76a03.s1 NCL_CGAP_Ew1 Homo sapiens cDNA clone IMAGE:1288140 similar to contains Alu repetitive element; contains element MER35 repetitive element;
5829	18919	32102	1.16	1.1E-01	AF020827.1	NT	6 Homo sapiens diacylglycerol kinase 3 (DGK3) gene, exon 6
5868	18957	32146	0.96	1.1E-01	AL110885.1	NT	Boltylis chirensis strain T4 cDNA library under conditions of nitrogen deprivation
5904	18990	32180	0.73	1.1E-01	BF339519.1	EST_HUMAN	602039176F1 NCL_CGAP_Bm64 Homo sapiens cDNA clone IMAGE:4186818 5'
5904	18990	32181	0.73	1.1E-01	BF339519.1	EST_HUMAN	602039176F1 NCL_CGAP_Bm64 Homo sapiens cDNA clone IMAGE:4186818 5'
5935	19021	32216	1.67	1.1E-01	X68851.1	NT	S. pombe ste8 gene encoding protein kinase
5971	19056	32256	4.98	1.1E-01	M86533.1	NT	Providencia rettgeri penicillin G amidase gene
6142	19217	32446	1.63	1.1E-01	AJ007973.1	NT	Homo sapiens LGMD2B gene
6164	19239	32470	1.54	1.1E-01	BE769152.1	EST_HUMAN	PM3-FT0024-130600-004-f12 FT0024 Homo sapiens cDNA
6184	19259	32493	6	1.1E-01	AW853699.1	EST_HUMAN	RC3-CT0254-280889-011-e01 CT0254 Homo sapiens cDNA
6594	19624	32889	0.5	1.1E-01	AL163282.2	NT	Homo sapiens chromosome 21 segment HS21C082
6571	19631	32898	1.26	1.1E-01	AF035748.1	EST_HUMAN	Homo sapiens chromosome 21 segment HS21C082
6617	19675	32953	0.84	1.1E-01	AE216307.1	EST_HUMAN	AF035748 Human salivary gland cell line HSG Homo sapiens cDNA clone RL43
6782	19816	33096	6.18	1.1E-01	O68636	SWISSPROT	gg76406.x1 Soares_NFL_T_G8C_S1 Homo sapiens cDNA clone IMAGE:1841086 3'
6862	19915		2.81	1.1E-01	AF032622.1	NT	ACETYL-COENZYME A SYNTHETASE (ACETATE-COA LIGASE) (ACYL-ACTIVATING ENZYME)
6862	20187	33512	2.23	1.1E-01	11432372	NT	Homo sapiens syntrophin 4 binding protein UNC-18a (UNC-18a) mRNA, complete cds
7249	19984	33281	0.65	1.1E-01	AE002155.1	NT	Homo sapiens phosphatidylinositol glycan, class B (PIGB), mRNA
7249	19984	33282	0.65	1.1E-01	AE002155.1	NT	Ureaplasma urealyticum section 58 of 58 of the complete genome
7393	25939		0.94	1.1E-01	BF382758.1	EST_HUMAN	Ureaplasma urealyticum section 58 of 58 of the complete genome
7523	26883	33850	1.58	1.1E-01	AF000006.1	NT	601816624F1 NIH_MGC_56 Homo sapiens cDNA clone IMAGE:4050853 5'
7781	20734	34105	7.69	1.1E-01	BF684628.1	EST_HUMAN	Pyrococcus horikoshii OTS genomic DNA, 1168001-1485000 nt, position (877)
7781	20734	34106	7.69	1.1E-01	BF684628.1	EST_HUMAN	602140976F1 NIH_MGC_46 Homo sapiens cDNA clone IMAGE:4302019 5'
7839	20786	34161	0.54	1.1E-01	AA985908.1	EST_HUMAN	602140976F1 NIH_MGC_46 Homo sapiens cDNA clone IMAGE:4302019 5'
7917	20860	34249	1.77	1.1E-01	P41087	SWISSPROT	cu44g03.s1 Soares_NFL_T_G8C_S1 Homo sapiens cDNA clone IMAGE:1829172 3'
7959	20900		0.79	1.1E-01	Z14098.1	NT	TRAB PROTEIN
							B. subtilis gene encoding hypothetical polyketide synthase

Table 4

Single Exon Probes Expressed in Bone Marrow

Probe SEQ ID NO.	Exon SEQ ID NO.	ORF SEQ ID NO.	Expression Signal	Most Similar (Top) Hit BLAST E Value	Top Hit Accession No.	Top Hit Database Source	Top Hit Descriptor
7861	20902	34294	2.89	1.1E-01	AA788784.1	EST_HUMAN	af31b08.s1 Soares parathyroid tumor NB-IPA Homo sapiens cDNA clone 1240403 3' similar to gb:J03483
8111	21048	34448	0.65	1.1E-01	BE782280.1	EST_HUMAN	CHROMOGRANIN A PRECURSOR (HUMAN);
8301	21270	34682	0.48	1.1E-01	U87482.1	NT	601470055F1 NIH_MGC_87 Homo sapiens cDNA clone IMAGE:3873229 5'
8550	21518	34836	1.61	1.1E-01	AA493574.1	EST_HUMAN	Methanococcus jannaschii section 34 of 150 of the complete genome
8550	21518	34837	1.61	1.1E-01	AA493574.1	EST_HUMAN	rh04g10.s1 NCI CGAP_Thy1 Homo sapiens cDNA clone IMAGE:943362
8588	21588	34882	1.1	1.1E-01	X91233.1	NT	rh04g10.s1 NCI CGAP_Thy1 Homo sapiens cDNA clone IMAGE:943362
8837	21605		1.06	1.1E-01	AW817918.1	EST_HUMAN	H.sapiens IL15 gene
8894	21682	35086	1.59	1.1E-01	AL134349.1	EST_HUMAN	PM1-ST0270-080200-001-408 S10270 Homo sapiens cDNA
9170	22136	35682	1.82	1.1E-01	U02482.1	NT	DKFZp547P194_r1 547 (synonym: hfbt1) Homo sapiens cDNA clone DKFZp547P194 5'
9283	22229	35680	0.98	1.1E-01	AJ807474.1	EST_HUMAN	Pedococcus acidiacetic H plasmid pSMB74 pediocin ACh production (pap) gene cluster papA, papB, papC and papD genes, complete cds
9362	22327	35756	0.56	1.1E-01	AF050081.1	NT	wf48c01.x1 Soares_NFL_T_GBC_S1 Homo sapiens cDNA clone IMAGE:2358816 3' similar to contains Alu repetitive element
9398	22363	35794	2.27	1.1E-01	AA192183.1	EST_HUMAN	Homo sapiens C16orf3 large protein mRNA, complete cds
9398	22363	35795	2.27	1.1E-01	AA192183.1	EST_HUMAN	zp83b12.r1 Stratagene muscle 837209 Homo sapiens cDNA clone IMAGE:627743 5'
9490	22454	35894	0.77	1.1E-01	Y12727.1	NT	zp83b12.r1 Stratagene muscle 837209 Homo sapiens cDNA clone IMAGE:627743 5'
9520	22483	35928	2.1	1.1E-01	T72675.1	EST_HUMAN	P.furiosus partial dph5 gene and argF gene
9546	22509		0.62	1.1E-01	BE883260.1	EST_HUMAN	yd18h03.s1 Soares fetal liver spleen 1NFLS Homo sapiens cDNA clone IMAGE:108725 3' similar to gb:M81181 SODIUM/POTASSIUM-TRANSPORTING ATPASE BETA-2 (HUMAN);
9777	22718		0.89	1.1E-01	BE142305.1	EST_HUMAN	601436872F1 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:3922048 5'
9853	22789		2.05	1.1E-01	BF085148.1	EST_HUMAN	CNA3-HT0142-271089-028-g11 HT0142 Homo sapiens cDNA
10270	23195		0.98	1.1E-01	AL161543.2	NT	MR2-GN0027-040900-005-a08 GN0027 Homo sapiens cDNA
10478	23398		0.45	1.1E-01	BE315509.1	EST_HUMAN	Arabidopsis thaliana DNA chromosome 4, contig fragment No. 43
10568	23488		1.01	1.1E-01	R80590.1	EST_HUMAN	601140231F1 NIH_MGC_9 Homo sapiens cDNA clone IMAGE:3049543 5'
10701	23623	37119	1.05	1.1E-01	U60528.1	NT	y83a09.s1 Soares placenta NB2H-P Homo sapiens cDNA clone IMAGE:147094 3'
11156	16103	29017	1.7	1.1E-01	F03265.1	EST_HUMAN	Ceratilis capitata yoyo retrotransposon gag-like, pol-like and env-like genes, complete cds
11267	24219		3.13	1.1E-01	AF168032.1	NT	HSC1RF022 normalized infant brain cDNA Homo sapiens cDNA clone c-1rf02 3'
11384	24331	37860	3.51	1.1E-01	R23708.1	EST_HUMAN	Carassius auratus activin beta A precursor, mRNA, complete cds
11392	24336	37868	1.54	1.1E-01	6981351	NT	yh35f12.r1 Soares placenta NB2H-P Homo sapiens cDNA clone IMAGE:131759 5' similar to contains Alu repetitive element; contains TAR1 repetitive element;
11542	24483	38035	2.18	1.1E-01	Z11910.1	NT	Rattus norvegicus Phosphofructokinase, liver, B-type (Pfkf), mRNA
11542	24483	38036	2.18	1.1E-01	Z11910.1	NT	Z.mobilis tgf and lig genes encoding tRNA guanine transglycosylase and DNA ligase
11638	24573	38137	3.66	1.1E-01	P17437	SWISSPROT	Z.mobilis tgf and lig genes encoding tRNA guanine transglycosylase and DNA ligase
							SKIN SECRETORY PROTEIN XP2 PRECURSOR (APEG PROTEIN)

Table 4

Single Exon Probes Expressed in Bone Marrow

Probe SEQ ID NO:	Exon SEQ ID NO:	ORF SEQ ID NO:	Expression Signal	Most Similar (Top) Hit BLAST E Value	Top Hit Accession No.	Top Hit Database Source	Top Hit Descriptor
12280	25084		2.03	1.1E-01	AA192153.1	EST_HUMAN	zp83b12.1 Stratagene muscle 637209 Homo sapiens cDNA clone IMAGE:827743 5'
12378	25161		3.68	1.1E-01	BE787023.1	EST_HUMAN	RC2-NT01112-120600-014-f03 NT01112 Homo sapiens cDNA
12626	25732		2.18	1.1E-01	BE974568.1	EST_HUMAN	601680551R2 NIH_MGC_83 Homo sapiens cDNA clone IMAGE:3950604 3'
13038	25580	31669	1.89	1.1E-01	BF238763.1	EST_HUMAN	601906350F1 NIH_MGC_54 Homo sapiens cDNA clone IMAGE:4134085 5'
13098	25812		1.32	1.1E-01	P14400	SWISSPROT	ELECTROMOTOR NEURON-ASSOCIATED PROTEIN 1
1206	14245		2.05	1.0E-01	O62855	SWISSPROT	DEOXYRIBONUCLEASE II PRECURSOR (DNASE II) (ACID DNASE) (LYSOSOMAL DNASE II)
1277	14312	27273	1.95	1.0E-01	AI985498.1	EST_HUMAN	we08401.x1 NC1_CGAP_Kd111 Homo sapiens cDNA clone IMAGE:2496577 3' similar to contains MER7.8
1393	14427	27396	2.26	1.0E-01	AL161504.2	NT	MER7 repetitive element;
2497	15500	28528	1.16	1.0E-01	AW451365.1	EST_HUMAN	Arabidopsis thaliana DNA chromosome 4, contig fragment No. 16
3524	16670	29483	1.32	1.0E-01	BF033991.1	EST_HUMAN	UH-B13-ab-4-07-0-U1.s1 NC1_CGAP_Sub5 Homo sapiens cDNA clone IMAGE:2738420 3'
3732	16774	29688	0.68	1.0E-01	BF238818.1	EST_HUMAN	601456301F1 NIH_MGC_68 Homo sapiens cDNA clone IMAGE:3859849 5'
3849	16889	29792	2.49	1.0E-01	AF297061.1	EST_HUMAN	601906489F1 NIH_MGC_54 Homo sapiens cDNA clone IMAGE:4134071 5'
3849	16889	29793	2.49	1.0E-01	AF297061.1	NT	Escherichia coli enterotoxin EspC (espC) gene, complete cds; and unknown genes
3978	17018	29830	2.88	1.0E-01	BF365703.1	EST_HUMAN	Escherichia coli enterotoxin EspC (espC) gene, complete cds; and unknown genes
4585	17607		0.88	1.0E-01	AF792349.1	EST_HUMAN	QV2-NT0048-160800-316-e05 NT0048 Homo sapiens cDNA
4741	17761	30655	1.32	1.0E-01	U50480.1	NT	sn32e04.y6 Gessler Wilms tumor Homo sapiens cDNA clone IMAGE:1700358 5'
4957	17872	30883	2.34	1.0E-01	AW952344.1	EST_HUMAN	Drosophila melanogaster tyrosine kinase p45 isoform (tir) mRNA, complete cds
5267	18274	31138	0.97	1.0E-01	AV721471.1	EST_HUMAN	EST3384414 IMAGE resequences, MAG8 Homo sapiens cDNA
5273	18279		1.04	1.0E-01	AV763960.1	EST_HUMAN	AV721471 HTB Homo sapiens cDNA clone HTBBQE10 5'
5394	18497		8.1	1.0E-01	W86490.1	EST_HUMAN	AV763960 MDS Homo sapiens cDNA clone MDSBQB11 5'
5492	18592		0.59	1.0E-01	X54015.1	NT	zh62h04.s1 Soares_fetal_liver_spleen_1NFLS_ST Homo sapiens cDNA clone IMAGE:416888 3'
5980	19083		0.87	1.0E-01	AK024472.1	NT	Xcampestris genes for sensor and regulator protein
6140	18218	32445	12.15	1.0E-01	AF274875.1	NT	Homo sapiens mRNA for FLJ00065 protein, partial cds
6468	19534	32782	0.9	1.0E-01	AA481879.1	EST_HUMAN	Homo sapiens growth factor receptor-bound protein 7 (GRB7) gene, complete cds
6483	19548	32797	0.65	1.0E-01	AA400039.1	EST_HUMAN	zwd41g10.s1 Soares ovary tumor NbHOT Homo sapiens cDNA clone IMAGE:756258 3' similar to contains
7220	20242		1.62	1.0E-01	R23821.1	EST_HUMAN	L1.13 L1 repetitive element;
8008	20847		2.16	1.0E-01	Y12488.1	NT	zwd7c12.s1 Soares_testis_NHT Homo sapiens cDNA clone IMAGE:743062 3'
8108	21045	34444	0.65	1.0E-01	AJ011400.1	NT	yh34h06.r1 Soares placenta Nb2HP Homo sapiens cDNA clone IMAGE:131675 5' similar to contains Alu
8108	21045	34445	0.65	1.0E-01	AJ011400.1	NT	repetitive element;
8265	21234	34645	0.63	1.0E-01	AA861091.1	EST_HUMAN	M.musculus wtn gene
						NT	Bos taurus mRNA for b17.2 subunit of NADH:ubiquinone oxidoreductase complex (complex I)
						NT	Bos taurus mRNA for b17.2 subunit of NADH:ubiquinone oxidoreductase complex (complex I)
						NT	ak32g01.s1 Soares_testis_NHT Homo sapiens cDNA clone IMAGE:1407888 3' similar to gb:M34182 CAMP-DEPENDENT PROTEIN KINASE, GAMMA-CATALYTIC SUBUNIT (HUMAN);

Table 4

Single Exon Probes Expressed in Bone Marrow

Probe SEQ ID NO:	Exon SEQ ID NO:	ORF SEQ ID NO:	Expression Signal	Most Similar (Top) Hit BLAST E Value	Top Hit Accession No.	Top Hit Database Source	Top Hit Descriptor
8287	21268	34685	0.45	1.0E-01	AF260225.1	NT	Homo sapiens TESTIN 2 and TESTIN 3 genes, complete cds, alternatively spliced
8287	21269	34686	0.45	1.0E-01	AF260225.1	NT	Homo sapiens TESTIN 2 and TESTIN 3 genes, complete cds, alternatively spliced
8506	21474		0.68	1.0E-01	4788365	NT	Homo sapiens fibroblast growth factor 13 (FGF13) mRNA
8837	21804		0.94	1.0E-01	AW188787.1	EST_HUMAN	x09601.x1 NCL CGAP_U4 Homo sapiens cDNA clone IMAGE:2875689 3' similar to gb-X17208 40S
9540	22503	35952	1.19	1.0E-01	AF102855.2	NT	RIBOSOMAL PROTEIN S4 (HUMAN); contains TAR1.83 TAR1 repetitive element;
9852	22788	36241	0.54	1.0E-01	R44983.1	EST_HUMAN	Rattus norvegicus synaptic SAPAP-interacting protein Synapton mRNA, complete cds
9865	22801		2.05	1.0E-01	M76728.1	NT	Yp33M04.s1 Soares Infant brain 1N1B Homo sapiens cDNA clone IMAGE:34549 3'
9908	22728		2.65	1.0E-01	AE001501.1	NT	Human pro-alpha-1 (V) collagen mRNA, complete cds
9922	22806	36259	0.61	1.0E-01	W01955.1	EST_HUMAN	Helicobacter pylori, strain J99 section 82 of 132 of the complete genome
10180	23105	36586	1.63	1.0E-01	BF240154.1	EST_HUMAN	z056c10.s1 Soares_fetal_heart_NbH19W Homo sapiens cDNA clone IMAGE:327282 3'
10285	23220	36703	9.1	1.0E-01	AB046798.1	NT	601805681F1 NIH_MGC_64 Homo sapiens cDNA clone IMAGE:4133487 5'
10295	23220	36704	9.1	1.0E-01	AB046798.1	NT	Homo sapiens mRNA for KIAA1579 protein, partial cds
10502	23424		0.95	1.0E-01	AW957425.1	EST_HUMAN	Homo sapiens mRNA for KIAA1579 protein, partial cds
10507	23429	36828	0.55	1.0E-01	T51952.1	EST_HUMAN	EST368615 MAGE resequences, MAGE Homo sapiens cDNA
10694	23616	37110	0.88	1.0E-01	BE792760.1	EST_HUMAN	y028a08.s1 Straigene fetal spleen (#837205) Homo sapiens cDNA clone IMAGE:72562 3' similar to contains Alu repetitive element
11018	23983		1.65	1.0E-01	AU159127.1	EST_HUMAN	601584804F1 NIH_MGC_7 Homo sapiens cDNA clone IMAGE:3939098 5'
11370	24317	37844	2.25	1.0E-01	BF242946.1	EST_HUMAN	AU159127 THYRO1 Homo sapiens cDNA clone THYRO1000885 3'
11370	24317	37845	2.25	1.0E-01	BF242946.1	EST_HUMAN	601877703F1 NIH_MGC_55 Homo sapiens cDNA clone IMAGE:4106089 5'
11728	24614	38191	3.16	1.0E-01	BE790543.1	EST_HUMAN	601877703F1 NIH_MGC_55 Homo sapiens cDNA clone IMAGE:4106089 5'
12365	25481		4.11	1.0E-01	BE537719.1	EST_HUMAN	601582558F1 NIH_MGC_7 Homo sapiens cDNA clone IMAGE:3939734 5'
12588	25282		2.18	1.0E-01	7682165	NT	601065554F1 NIH_MGC_10 Homo sapiens cDNA clone IMAGE:3451833 5'
12604	25305		2.58	1.0E-01	X00854.1	NT	Homo sapiens KIAA0514 gene product (KIAA0514), mRNA
12786	25425		1.45	1.0E-01	AA737881.1	EST_HUMAN	Drosophila melanogaster ftz gene
12885	25481		4.47	1.0E-01	BE537719.1	EST_HUMAN	mx11c08.s1 NCI CGAP GC3 Homo sapiens cDNA clone IMAGE:1255780 3'
12943	25522		1.32	1.0E-01	BE158905.1	EST_HUMAN	601065554F1 NIH_MGC_10 Homo sapiens cDNA clone IMAGE:3451833 5'
12959	25581		6.14	1.0E-01	U66834.1	NT	QV4-HT0401-211289-064-g03 HT0401 Homo sapiens cDNA
13016	25568		7.59	1.0E-01	AP001507.1	NT	Saccharomyces cerevisiae suppressor of ABF1 (SAB2) gene, complete cds
2789	15781	28797	1.74	9.9E-02	AF274008.1	NT	Bacillus halodurans genomic DNA, section 1/14
2788	15790	28808	1.71	9.9E-02	BE545554.1	EST_HUMAN	Drosophila melanogaster cAMP-dependent protein kinase type II regulatory subunit (pka-R1) mRNA, complete cds
2788	15790	28809	1.71	9.9E-02	BE545554.1	EST_HUMAN	601070219F1 NIH_MGC_12 Homo sapiens cDNA clone IMAGE:3456385 5'
3280	16334	28254	1.48	9.9E-02	AF089810.1	NT	601070219F1 NIH_MGC_12 Homo sapiens cDNA clone IMAGE:3456385 5'

Table 4

Single Exon Probes Expressed in Bone Marrow

Probe SEQ ID NO:	Exon SEQ ID NO:	ORF SEQ ID NO:	Expression Signal	Most Similar (Top) Hit BLAST E Value	Top Hit Accession No.	Top Hit Database Source	Top Hit Descriptor
3974	17014	28928	0.97	9.9E-02	AI821637.1	EST_HUMAN	zu45cd03.x5 Soares ovary tumor NIDHOT Homo sapiens cDNA clone IMAGE:740832 3'
4707	17728	30622	1.02	9.9E-02	BE674249.1	EST_HUMAN	7d77c12.x1 NC1_CGAP_Lu24 Homo sapiens cDNA clone IMAGE:3278998 3'
7044	20066	33373	2.59	9.9E-02	BE613498.1	EST_HUMAN	601504252F1 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:3906094 5'
7163	18394	31239	7.77	9.9E-02	D83710.1	NT	Aspergillus terreus BSD mRNA for blesticidin S deaminase, complete cds
8247	21216	34824	0.65	9.9E-02	AW103088.1	EST_HUMAN	xd43cd08.x1 NC1_CGAP_Ov23 Homo sapiens cDNA clone IMAGE:2596528 3' similar to contains Alu repetitive element; contains element MIR MIR repetitive element ;
8247	21216	34826	0.66	9.9E-02	AW103088.1	EST_HUMAN	xd43cd08.x1 NC1_CGAP_Ov23 Homo sapiens cDNA clone IMAGE:2596528 3' similar to contains Alu repetitive element; contains element MIR MIR repetitive element ;
9912	22616	36068	1.23	9.9E-02	6755111	NT	Mus musculus phospholipid transfer protein (Pltp), mRNA
565	13635		1.88	9.8E-02	X56338.1	NT	O. sativa RAMY6C gene for alpha-amylase
3160	16216	29130	4.25	9.8E-02	AF184274.1	NT	Daucus carota lacunarothocyanidin dioxygenase 2 (LDOX) mRNA, LDOX-2 allele, complete cds
4251	17280	30160	6.93	9.8E-02	AF257328.1	NT	Leptospira maculans beta-tubulin mRNA, complete cds
4251	17280	30161	6.93	9.8E-02	AF257328.1	NT	Leptospira maculans beta-tubulin mRNA, complete cds
7723	20679		0.94	9.8E-02	X54133.1	NT	Human HPTP delta mRNA for protein tyrosine phosphatase delta
9609	22613		1.21	9.8E-02	M61943.1	NT	Human laminin B1 chain gene, exon 28
11788	23943	37465	1.83	9.8E-02	BF037421.1	EST_HUMAN	601460783F1 NIH_MGC_60 Homo sapiens cDNA clone IMAGE:3864287 5'
12330	25131		1.84	9.8E-02	8393761	NT	Rattus norvegicus microtubule-associated protein tau (Mept), mRNA
1352	14387	27357	1.75	9.7E-02	AB005808.1	NT	Aloe arborescens mRNA for NADP-malic enzyme, complete cds
1589	14621		1.33	9.7E-02	4503710	NT	Homo sapiens fibroblast growth factor receptor 3 (echondroplasia, thanatophoric dwarfism) (FGFR3) mRNA
2270	15283	28309	2.78	9.7E-02	BE168680.1	EST_HUMAN	QV1-HT0516-070300-095-ab4 HT0516 Homo sapiens cDNA
4008	17047		4.89	9.7E-02	Q99795	SWISSPROT	CELL SURFACE A33 ANTIGEN PRECURSOR (GLYCOPROTEIN A33)
5418	18521	31398	0.88	9.7E-02	AF039189.1	NT	Caulobacter crescentus thymidylate kinase (tnk) and DNA polymerase III delta prime subunit (dnaC) genes, complete cds
5418	18521	31399	0.88	9.7E-02	AF039189.1	NT	Caulobacter crescentus thymidylate kinase (tnk) and DNA polymerase III delta prime subunit (dnaC) genes, complete cds
6130	19207	32432	1.48	9.7E-02	AW954478.1	EST_HUMAN	EST368546 MAGC resequences, MAGC Homo sapiens cDNA
7517	20482	33943	3.36	9.7E-02	Z99119.1	NT	Bacillus subtilis complete genome (section 16 of 21); from 2987771 to 3213410
8315	21284	34697	1.09	9.7E-02	N22788.1	EST_HUMAN	yw41cd03.s1 Wetzmann Olfactory Epithelium Homo sapiens cDNA clone IMAGE:254788 3'
8315	21284	34698	1.09	9.7E-02	N22788.1	EST_HUMAN	yw41cd03.s1 Wetzmann Olfactory Epithelium Homo sapiens cDNA clone IMAGE:254788 3'
9201	22167	35597	1.52	9.7E-02	AB53084.1	EST_HUMAN	w078b06.x1 NC1_CGAP_Ov38 Homo sapiens cDNA clone IMAGE:2549747 3' similar to gb:U52851_rna1
11629	24470		2.01	9.7E-02	U68337.1	NT	PEPTIDYL-PROLYL CIS-TRANS ISOMERASE A (HUMAN); Mus musculus ligatin (Lgtn) mRNA, partial cds
2032	15051	28066	1.27	9.6E-02	AI080721.1	EST_HUMAN	oz47d11.x1 Soares_NHMPu_S1 Homo sapiens cDNA clone IMAGE:1679485 3'

Table 4

Single Exon Probes Expressed in Bone Marrow

Probe SEQ ID NO:	Exon SEQ ID NO:	ORF SEQ ID NO:	Expression Signal	Most Similar (Top) Hit BLAST E Value	Top Hit Accession No.	Top Hit Database Source	Top Hit Descriptor
2032	15051	28087	1.27	9.6E-02	AI080721.1	EST_HUMAN	oz47d11.x1 Soares NIHIMPu_S1 Homo sapiens cDNA clone IMAGE:1678485 3'
4372	17389	30279	6.7	9.6E-02	Z32686.2	NT	Proteus mirabilis fimbrial operon, strain H4320
5037	18050	30830	1.16	9.6E-02	AW968230.1	EST_HUMAN	EST379303 IMAGE resequences, MAGI Homo sapiens cDNA
6225	19299		2.83	9.6E-02	BE910039.1	EST_HUMAN	601498088F1 NIH_MGC_70 Homo sapiens cDNA clone IMAGE:3900165 5'
8719	21687		0.65	9.6E-02	AI137084.1	EST_HUMAN	AU137084 PLACE1 Homo sapiens cDNA clone PLACE1005740 5'
9902	22854	36315	1.35	9.6E-02	AV687688.1	EST_HUMAN	AV687688 GKC Homo sapiens cDNA clone GKCAAH02 5'
10231	23156		1.29	9.6E-02	BE894895.1	EST_HUMAN	601434030F1 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:3916363 5'
10389	23321	36806	1.27	9.6E-02	AJ243211.1	NT	Homo sapiens DMBT1 candidate tumour suppressor gene, exons 1 to 55
10389	23321	36808	1.27	9.6E-02	AJ243211.1	NT	Homo sapiens DMBT1 candidate tumour suppressor gene, exons 1 to 55
10479	23401	36808	0.44	9.6E-02	BF677270.1	EST_HUMAN	602086798F1 NIH_MGC_83 Homo sapiens cDNA clone IMAGE:4260869 5'
10511	23433	36830	1.37	9.6E-02	AB013985.1	NT	Anthrimum majus transposon Tam3 pseudogene for transposase (in S-6 copy)
10511	23433	36831	1.37	9.6E-02	AB013985.1	NT	Anthrimum majus transposon Tam3 pseudogene for transposase (in S-6 copy)
10621	23543	37043	3.5	9.6E-02	P08174	SWISSPROT	COMPLEMENT DECAY-ACCELERATING FACTOR PRECURSOR (CD55)
11094	24054	37577	5.31	9.6E-02	Z79702.1	NT	Mycobacterium tuberculosis H37Rv complete genome, segment 102/162
12028	24904	38498	1.53	9.6E-02	AA625755.1	EST_HUMAN	zu91g01.s1 Soares testis NIH Homo sapiens cDNA clone IMAGE:745392 3'
12933	25515		1.81	9.6E-02	H14599.1	EST_HUMAN	ym19H03.s1 Soares infant brain NIH Homo sapiens cDNA clone IMAGE:48653 3'
12996	25563	31719	1.51	9.6E-02	BE728219.1	EST_HUMAN	601563355F1 NIH_MGC_20 Homo sapiens cDNA clone IMAGE:3832808 5'
4128	17161	30050	2.68	9.5E-02	AW992395.1	EST_HUMAN	CM2-BN0023-050200-087-412 BN0023 Homo sapiens cDNA
5748	18842	32025	0.81	9.5E-02	P51854	SWISSPROT	TRANSETOLASE 2 (TK 2) (TRANSETOLASE RELATED PROTEIN)
7294	20266	33601	0.51	9.5E-02	AA780728.1	EST_HUMAN	ac88a09.s1 Stratagene fetal retina 937202 Homo sapiens cDNA clone IMAGE:967736 3'
7522	20487	33949	4.14	9.5E-02	AB003473.1	NT	Trimeresurus flavoviridis DNA for phosphodiesterase A2 inhibitor, complete cds
7818	20767	34144	7.10	9.5E-02	AL101538.2	NT	Arabidopsis thaliana DNA chromosome 4, contig fragment No. 38
7963	18842	32025	0.89	9.5E-02	P51854	SWISSPROT	TRANSETOLASE 2 (TK 2) (TRANSETOLASE RELATED PROTEIN)
8212	21181	34591	2.85	9.5E-02	BF035961.1	EST_HUMAN	601453842F1 NIH_MGC_68 Homo sapiens cDNA clone IMAGE:3857243 5'
8212	21181	34592	2.85	9.5E-02	BF035961.1	EST_HUMAN	601453842F1 NIH_MGC_68 Homo sapiens cDNA clone IMAGE:3857243 5'
11039	24003	37528	2.31	9.5E-02	BF035961.1	EST_HUMAN	601453842F1 NIH_MGC_68 Homo sapiens cDNA clone IMAGE:3857243 5'
11039	24003	37529	2.31	9.5E-02	BF035961.1	EST_HUMAN	601453842F1 NIH_MGC_68 Homo sapiens cDNA clone IMAGE:3857243 5'
1850	14878	27871	4.07	9.4E-02	BF671063.1	EST_HUMAN	602150882F1 NIH_MGC_81 Homo sapiens cDNA clone IMAGE:4281917 5'
3894	16834	29844	4.91	9.4E-02	Z330059.1	NT	M.leptotrichum DNA for CONTIG MC073
5240	18248		0.96	9.4E-02	X98108.1	NT	Lactobacillus bacteriophage phig1e complete genomic DNA
6450	19515	32768	1.21	9.4E-02	AF087363.1	NT	Triticum aestivum heat shock protein 101 (Hsp101a) mRNA, complete cds
7851	20797	34173	0.54	9.4E-02	L78833.1	NT	Human BRCA1, Rho7 and val genes, complete cds, and p35 gene, partial cds
8947	21913		2.5	9.4E-02	Z49863.1	NT	Acinetobacter sp. cysD, cdbQ, sodM, lysS, rubA, rubB, estB, oxvR, ppk, mtgA, ORF2 and ORF3 genes

Table 4

Single Exon Probes Expressed In Bone Marrow

Probe SEQ ID NO:	Exon SEQ ID NO:	ORF SEQ ID NO:	Expression Signal	Most Similar (Top) Hit BLAST E Value	Top Hit Accession No.	Top Hit Database Source	Top Hit Descriptor
11280	20797	34173	2.22	9.4E-02	L78833.1	NT	Human BRCA1, Rho7 and vail genes, complete cds, and lpf35 gene, partial cds
12212	25828		10.73	9.4E-02	U31815.1	NT	Rattus norvegicus calcium channel alpha-1C subunit (ROB2) mRNA, partial cds
13087	25817	31682	1.37	9.4E-02	U27699.1	NT	Human pefBGT-1 betaine-GABA transporter mRNA, complete cds
3002	16060		1.71	9.3E-02	4808280	NT	Human sapiens BA11-associated protein 3 (BAIAP3) mRNA
3041	16099		7.39	9.3E-02	6912526	NT	Human sapiens nescopharyngeal epithelium specific protein 1 (NESG1), mRNA
3270	16324	28247	2.01	9.3E-02	BF57651.1	EST_HUMAN	602133089F1 NIH_MGC_81 Homo sapiens cDNA clone IMAGE:4288269 5'
4179	17210	30098	3.73	9.3E-02	BE391943.1	EST_HUMAN	601286082F1 NIH_MGC_44 Homo sapiens cDNA clone IMAGE:3607653 5'
4179	17210	30097	3.73	9.3E-02	BE391943.1	EST_HUMAN	601286082F1 NIH_MGC_44 Homo sapiens cDNA clone IMAGE:3607653 5'
4768	17788		2.29	9.3E-02	AV732224.1	EST_HUMAN	AV732224 HTF Homo sapiens cDNA clone HTFAUA06 5'
5745	18839		0.66	9.3E-02	AP001507.1	NT	Bacillus halodurans genomic DNA, section 1/4
8163	21091	34490	0.59	9.3E-02	AL163210.2	NT	Homo sapiens chromosome 21 segment HS21C010
8591	21559	34976	0.6	9.3E-02	AW568007.1	EST_HUMAN	EST89 Human Fetal Brain MATCHMAKER cDNA Library Homo sapiens cDNA
9480	22444		0.42	9.3E-02	AL113179.1	NT	Botrytis cinerea strain T4 cDNA library under conditions of nitrogen deprivation
10068	22995	36464	2.32	9.3E-02	BE362831.2	EST_HUMAN	601665968F1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:3855981 3'
10651	23473	36967	3.65	9.3E-02	Q15034	SWISSPROT	HYPOTHETICAL PROTEIN KIAA0032
10551	23473	36968	3.65	9.3E-02	Q15034	SWISSPROT	HYPOTHETICAL PROTEIN KIAA0032
10883	23605		3.59	9.3E-02	AW206117.1	EST_HUMAN	UHH-B11-adv-h-05-0-U1.s1 NCI_CGAP_Sub3 Homo sapiens cDNA clone IMAGE:2723553 3'
12482	25750		2.55	9.3E-02	AJ249850.1	NT	Photobacterium damselae subsp. damsela partial gyrB gene for DNA gyrase B subunit
12831	25778		9.12	9.3E-02	AW469850.1	EST_HUMAN	Hd28h12x1 Soares_NFL_T_GBC S1 Homo sapiens cDNA clone IMAGE:2910887 3'
13040	25824		2.24	9.3E-02	AF100958.1	NT	Mus musculus major histocompatibility locus class II region; Fes-binding protein Daxx (DAXX) gene, partial cds; Bing1 (BING1), tapasin (tapasin), RafGDS-like factor (RLF), KE2 (KE2), BING4 (BING4), beta1, 3-galactosyl transferase (beta1,3-galactosyl tr>
231	13331	26252	7.81	9.2E-02	U60315.1	NT	Molluscum contagiosum virus subtype 1, complete genome
231	13331	26253	7.81	9.2E-02	U60315.1	NT	Molluscum contagiosum virus subtype 1, complete genome
231	13331	26254	7.81	9.2E-02	U60315.1	NT	Molluscum contagiosum virus subtype 1, complete genome
2236	15280		2.72	9.2E-02	R54156.1	EST_HUMAN	y98807.r1 Soares Infant brain 1NIB Homo sapiens cDNA clone IMAGE:41618 5'
3194	16249	29167	3.72	9.2E-02	Q29831	SWISSPROT	MAJOR EPIDIDYMIS-SPECIFIC PROTEIN E4 (EPIDIDYMAL PROTEIN BE-20)
3316	16369	28289	0.99	9.2E-02	AA634554.1	EST_HUMAN	nt78001.s1 NCI_CGAP_Cos8 Homo sapiens cDNA clone IMAGE:926136 3'
3699	16844		1.27	9.2E-02	6755215	NT	Mus musculus pre T-cell antigen receptor alpha (Pctra), mRNA
4268	17295		1.24	9.2E-02	U92048.1	NT	Human herpesvirus 1 strain KOS-63, latency-associated transcript, promoter region
4337	17364		0.94	9.2E-02	BE299722.1	EST_HUMAN	600844365F1 NIH_MGC_17 Homo sapiens cDNA clone IMAGE:2960176 5'
4668	17699	30575	1.4	9.2E-02	X68402.1	NT	G.gallus Mta-CK gene
8342	21311	34725	1.87	9.2E-02	T49920.1	EST_HUMAN	y889-c09.r1 Stratiotes placenta (#937225) Homo sapiens cDNA clone IMAGE:69808 5' similar to similar to gb:X59309 GUANINE NUCLEOTIDE-BINDING PROTEIN G(S), ALPHA SUBUNIT (HUMAN)

Page 124 of 546
Table 4

Single Exon Probes Expressed in Bone Marrow

Probe SEQ ID NO:	Exon SEQ ID NO:	ORF SEQ ID NO:	Expression Signal	Most Similar (Top) Hit BLAST E Value	Top Hit Accession No.	Top Hit Database Source	Top Hit Descriptor
8517	21485	34899	2.19	9.2E-02	X85286.1	NT	H. vulgare xylose isomerase gene
12002	24879	38475	1.74	9.2E-02	AF026562.3	NT	Mesocricetus auratus oviductin precursor (OVI) gene, complete cds
13019	26988		1.31	9.2E-02	11466872	NT	Podospira anserina mitochondrion, complete genome
423	13118	28017	7.82	9.1E-02	X7086.1	NT	O. cuniculus K12 keratin gene
3884	16727		0.98	9.1E-02	AW372589.1	EST_HUMAN	PM2-BT0349-161289-001-02 BT0349 Homo sapiens cDNA
4510	17535	30419	1.5	9.1E-02	AL161554.2	NT	Arabidopsis thaliana DNA chromosome 4, contig fragment No. 54
5304	18307		0.98	9.1E-02	AB010898.1	NT	Clona Intestitralis endostyle-specific mRNA, complete cds
5819	18809	32063	1.27	9.1E-02	AF129758.1	NT	Homo sapiens MSH455 gene, partial cds; and CLIC1, DDAH, G8b, G8c, G5b, G6d, G6a, G6f, BAT5, G5b, CSK2B, BAT4, G4, Apo M, BAT3, BAT2, AIF-1, 1C7, LST-1, LTB, TNF, and LTA genes, complete cds
7528	28000		0.52	9.1E-02	AF028308.1	NT	Homo sapiens chromosome 9 duplication of the T cell receptor beta locus and trypsinogen gene families
7616	20576	33939	12.92	9.1E-02	AW160658.1	EST_HUMAN	au74605.y1 Schneider fetal brain 00004 Homo sapiens cDNA clone IMAGE:2781868 5
7937	20879	34269	0.74	9.1E-02	AF000061.1	NT	Aenopyrum pennix genome DNA, section 47
7977	20918	34307	0.87	9.1E-02	U39073.1	NT	Mus musculus thymopoietin zeta mRNA, complete cds
8278	22242	35671	0.85	9.1E-02	Y14378.1	NT	Homo sapiens gamma adducin gene, exon 9
10793	23714		1.84	9.1E-02	T02894.1	EST_HUMAN	FB19F10 Fetal brain, Strategene Homo sapiens cDNA clone FB19F10 3'end
10823	23744	37245	1.24	9.1E-02	S74058.1	NT	Tg616-Cy4 actin [Tritreussea gratillae=sea urchins, embryos, Genomic, 5275 nt]
10852	23772	37271	0.8	9.1E-02	Y11187.1	NT	A. thaliana RH1, TC1, G14587-5, G14587-6, and PRL1 genes
12151	25009		3.53	9.1E-02	8633494	NT	Bacteriophage Mu, complete genome
12393	25919		2.15	9.1E-02	AA179801.1	EST_HUMAN	zp38h12.s1 Strategene muscle 937209 Homo sapiens cDNA clone IMAGE:611783 3' similar to
12468	25217		1.82	9.1E-02	AF052885.1	NT	SW:TRT3_HUMAN P48378 TROPONIN T, FAST SKELETAL MUSCLE, ISOFORM BETA ;
12916	25771		1.78	9.1E-02	AJ291380.1	NT	Rattus norvegicus cell cycle protein p55CDC gene, complete cds
						NT	Homo sapiens partial MUC3B gene for MUC3B mucin, exons 1-11
							FOLATE RECEPTOR ALPHA PRECURSOR (FR-ALPHA) (FOLATE RECEPTOR 1) (FOLATE RECEPTOR, ADULT) (ADULT FOLATE-BINDING PROTEIN) (FBP) (OVARIAN TUMOR-ASSOCIATED ANTIGEN MOV18) (KB CELLS FBP)
746	19807	28747	6.38	9.0E-02	P15328	SWISSPROT	ANTIGEN MOV18 (KB CELLS FBP)
							h339g10.x1 NC1_CGAP_Lu24 Homo sapiens cDNA clone IMAGE:3175842 3' similar to contains Alu repetitive element;
1640	14672	27645	6.7	9.0E-02	BE220482.1	EST_HUMAN	HIV-1 p8c095-08 from USA envelope glycoprotein (env) gene, partial cds
2815	15807	28826	6.5	9.0E-02	AF138522.1	NT	HIV-1 p8c095-08 from USA envelope glycoprotein (env) gene, partial cds
2815	15807	28827	6.5	9.0E-02	AF138522.1	NT	Dicystostellum discoidium spore coat structural protein SP65 (cotE) gene, complete cds
3347	16398	28321	1.11	9.0E-02	AF279135.1	NT	corticosteroid-binding globulin [Saimiri sciureus=squirrel monkeys, liver, mRNA, 1474 nt]
4328	17358	30243	0.83	9.0E-02	S98757.1	NT	corticosteroid-binding globulin [Saimiri sciureus=squirrel monkeys, liver, mRNA, 1474 nt]
4328	17356	30244	0.83	9.0E-02	S98757.1	NT	corticosteroid-binding globulin [Saimiri sciureus=squirrel monkeys, liver, mRNA, 1474 nt]

Table 4

Single Exon Probes Expressed in Bone Marrow

Probe SEQ ID NO:	Exon SEQ ID NO:	ORF SEQ ID NO:	Expression Signal	Most Similar (Top) Hit BLAST E Value	Top Hit Accession No.	Top Hit Database Source	Top Hit Descriptor
4448	17474	30363	1.26	9.0E-02	P55288	SWISSPROT	LAMININ BETA-2 CHAIN PRECURSOR (S-LAMININ)
4695	17716	30611	2.37	9.0E-02	X65740.2	NT	Plasmodium falciparum P-type ATPase 3 gene
5261	18269	31137	0.93	9.0E-02	Q24597	SWISSPROT	REGULATORY PROTEIN ZESTIE
6110	19189	32409	14.12	9.0E-02	W56037.1	EST_HUMAN	za68a12.r1 Soares_fetal_jung_NbHL10W Homo sapiens cDNA clone IMAGE:287694 5' similar to PIR:S52171 S52171 small G protein - human ;
6880	19932		1.14	9.0E-02	BF062851.1	EST_HUMAN	7n03d003.x1 NC1_CGAP_Cot16 Homo sapiens cDNA clone IMAGE:3320845 3' similar to contains Alu repetitive element;
6831	20155	33473	0.86	9.0E-02	R62805.1	EST_HUMAN	yf11b08.s1 Soares placenta Nb2HP Homo sapiens cDNA clone IMAGE:138903 3'
12762	25404		2.09	9.0E-02	AF022238.1	NT	Escherichia coli strain E2348/69 pathogenicity island, rOrf1 (rOrf1), rOrf2 (rOrf2), EscR (escR), EscS (escS), EscT (escT), EscU (escU), CasD (casD), EscC (escC), EscJ (escJ), SepZ (sepZ), EscV (escV), EscN (escN), SepQ (sepQ), Tr (tr), OrfU (orfU), >
1432	14498	27442	2.04	8.9E-02	BF701593.1	EST_HUMAN	602129030F2 NIH_MGC_56 Homo sapiens cDNA clone IMAGE:4285951 5'
1432	14498	27443	2.04	8.9E-02	BF701593.1	EST_HUMAN	602129030F2 NIH_MGC_56 Homo sapiens cDNA clone IMAGE:4285951 5'
2368	16403	28429	1.1	8.9E-02	BE153572.1	EST_HUMAN	PM0-HT0339-261198-003-d01 HT0339 Homo sapiens cDNA
4227	17256		2	8.9E-02	AF286055.1	NT	Atrichum angustatum AtranFol2 protein (AtranFol2) gene, partial cds
5950	19036	32230	2.64	8.9E-02	AW452122.1	EST_HUMAN	UIH-B13-elo-f08-0-J1.s1 NC1_CGAP_Sub5 Homo sapiens cDNA clone IMAGE:3068294 3'
5950	19036	32231	2.64	8.9E-02	AW452122.1	EST_HUMAN	UIH-B13-elo-f08-0-J1.s1 NC1_CGAP_Sub5 Homo sapiens cDNA clone IMAGE:3068294 3'
5968	18051	32262	3.5	8.9E-02	11433478	NT	Homo sapiens similar to endoglycan (H. sapiens) (LOC63107), mRNA
7399	20367	33720	1.49	8.9E-02	P47259	SWISSPROT	FOLD BIFUNCTIONAL PROTEIN [INCLUDES: METHYLENETETRAHYDROFOLATE DEHYDROGENASE; METHYLENETETRAHYDROFOLATE CYCLOHYDROLASE]
7807	20757		2.02	8.9E-02	Z79021.1	NT	H. sapiens flow-sorted chromosome 6 HindIII fragment, SC8pA20F8
8385	21354	34762	0.77	8.9E-02	P29475	SWISSPROT	NITRIC-OXIDE SYNTHASE, BRAIN (NOS, TYPE I) (NEURONAL NOS) (N-NOS) (NNOS) (CONSTITUTIVE NOS) (NC-NOS) (BNOS)
8468	21437	34855	0.7	8.9E-02	BF701695.1	EST_HUMAN	602129111F2 NIH_MGC_56 Homo sapiens cDNA clone IMAGE:4285827 5'
8468	21437	34856	0.7	8.9E-02	BF701695.1	EST_HUMAN	602129111F2 NIH_MGC_56 Homo sapiens cDNA clone IMAGE:4285827 5'
8945	21911	35338	5.01	8.9E-02	AA309319.1	EST_HUMAN	EST180187 Liver, hepatocellular carcinoma Homo sapiens cDNA 5' end
9978	22905	36369	0.62	8.9E-02	A1285627.1	EST_HUMAN	qu55c05.x1 NC1_CGAP_Lym6 Homo sapiens cDNA clone IMAGE:1968680 3' similar to contains MER10.b1 MER10 repetitive element;
9978	22905	36370	0.62	8.9E-02	A1285627.1	EST_HUMAN	qu55c05.x1 NC1_CGAP_Lym6 Homo sapiens cDNA clone IMAGE:1968680 3' similar to contains MER10.b1 MER10 repetitive element;
10090	23016	36492	0.67	8.9E-02	AA339359.1	EST_HUMAN	EST44454 Fetal brain 1 Homo sapiens cDNA 5' end
12150	25743		1.3	8.9E-02	P30143	SWISSPROT	HYPOPHYSICAL 51.7 KD PROTEIN IN THRC-TALB INTERGENIC REGION (ORF8)
12211	25778		1.91	8.9E-02	P19624	SWISSPROT	MYOSIN-2 ISOFORM
12367	25154		9.08	8.9E-02	BF696918.1	EST_HUMAN	602129882F1 NIH_MGC_56 Homo sapiens cDNA clone IMAGE:4286180 5'

Table 4

Single Exon Probes Expressed in Bone Marrow

Probe SEQ ID NO:	Exon SEQ ID NO:	ORF SEQ ID NO:	Expression Signal	Most Similar (Top) Hit BLAST E Value	Top Hit Accession No.	Top Hit Database Source	Top Hit Descriptor
12530	25258		1.81	8.9E-02	6080220	NT	Mus musculus hippocampus abundant gene transcript 1 (Hiat1), mRNA
12571	25282		1.5	8.9E-02	U28895.1	NT	Human 4-hydroxyphenylpyruvate-dioxygenase gene, complete cds
1374	14408	27378	1.25	8.8E-02	Q27474	SWISSPROT	PROBABLE DNA LIGASE (POLYDEOXYRIBONUCLEOTIDE SYNTHASE [ATP])
3921	16981	29874	1.03	8.8E-02	AA289128.1	EST_HUMAN	EST11595 Uterus Homo sapiens cDNA 5' end
4066	17101		3.7	8.8E-02	O00288	SWISSPROT	TRANSCRIPTION INITIATION FACTOR TFIIID 135 KDA SUBUNIT (TAFII135) (TAFII130)
4287	17298		1.28	8.8E-02	4502804	NT	Homo sapiens chromogranin A (parathyroid secretory protein 1) (CHGA) mRNA
4931	17359		2.3	8.8E-02	4580423	NT	Homo sapiens paired box gene 6 (enitridia, keratitis) (PAX6), isoform b, mRNA
7792	20744		0.86	8.8E-02	D17520.1	NT	Sheep mRNA for angiotensinogen, complete cde
8339	22304	35732	1.32	8.8E-02	AA151872.1	EST_HUMAN	zn9a05.s1 Stragene cdon (#837204) Homo sapiens cDNA clone IMAGE:568288 3'
11453	24396	37941	3.43	8.8E-02	BE284455.1	EST_HUMAN	601191770F1 NIH_MGC_7 Homo sapiens cDNA clone IMAGE:3535948 5'
11453	24396	37942	3.43	8.8E-02	BE284455.1	EST_HUMAN	601191770F1 NIH_MGC_7 Homo sapiens cDNA clone IMAGE:3535948 5'
11593	24531	38088	5.78	8.8E-02	AL040129.1	EST_HUMAN	DKFZp434D1313_J1 434 (synonym: hhes3) Homo sapiens cDNA clone DKFZp434D1313 5'
12441	25204	31827	2.4	8.8E-02	Z71591.1	NT	S. cerevisiae chromosome XIV reading frame ORF YNL285w
1654	14686	27661	1.55	8.7E-02	A167281.1	EST_HUMAN	aa65b01.s1 Soares_NhlHMPu_S1 Homo sapiens cDNA clone IMAGE:1661161 3'
3707	16750	29665	3.84	8.7E-02	U82695.2	NT	Homo sapiens zinc finger protein 92 (ZFP92), expressed-Xq28STS protein (XQ28ORF), and biglycan (BGN) genes, complete cds; and plasma membrane calcium ATPase isoform 3 (PMCA3) gene, partial cds
3707	16750	29666	3.84	8.7E-02	U82695.2	NT	Homo sapiens zinc finger protein 92 (ZFP92), expressed-Xq28STS protein (XQ28ORF), and biglycan (BGN) genes, complete cds; and plasma membrane calcium ATPase isoform 3 (PMCA3) gene, partial cds
4736	17756	30650	1.2	8.7E-02	AF178636.1	NT	Mus musculus JNK interacting protein-3a (Jip3) mRNA, complete cds
5138	18147		1.02	8.7E-02	AE000895.1	NT	Methanobacterium thermoautotrophicum from bases 1176181 to 1189406 (section 101 of 148) of the complete genome
5166	18175	31053	0.97	8.7E-02	6563384	NT	Homo sapiens protein kinase C, η 1 (PRKCN), mRNA
5387	18490	31365	6.04	8.7E-02	AA286875.1	EST_HUMAN	zs55g08.s1 NCI_CGAP_GCB1 Homo sapiens cDNA clone IMAGE:701438 3'
5387	18490	31368	6.04	8.7E-02	AA286875.1	EST_HUMAN	zs55g08.s1 NCI_CGAP_GCB1 Homo sapiens cDNA clone IMAGE:701438 3'
7022	20147	33468	0.81	8.7E-02	AJ271855.2	NT	Mus musculus partial Kcnq1 gene for potassium channel protein, exons 10-14
7022	20147	33467	0.81	8.7E-02	AJ271855.2	NT	Mus musculus partial Kcnq1 gene for potassium channel protein, exons 10-14
7243	19878	33276	0.63	8.7E-02	AF281342.1	NT	Oncorhynchus mykiss TAT-binding protein 1 mRNA, partial cds
8890	21827	35250	0.71	8.7E-02	AE004767.1	NT	Pseudomonas aeruginosa PAO1, section 348 of 529 of the complete genome
8890	21827	35251	0.71	8.7E-02	AE004767.1	NT	Pseudomonas aeruginosa PAO1, section 348 of 529 of the complete genome
11057	24030		2.15	8.7E-02	L04758.1	NT	Oryctolagus cuniculus cytochrome P-450 (CYP4A4) gene, 5' end
11641	24578	38144	1.79	8.7E-02	AJ007763.1	NT	Glucobacter oxydans tRNA-Ile and tRNA-Ala genes

Page 127 of 546
Table 4

Single Exon Probes Expressed In Bone Marrow

Probe SEQ ID NO:	Exon SEQ ID NO:	ORF SEQ ID NO:	Expression Signal	Most Similar (Top) Hit BLAST E Value	Top Hit Accession No.	Top Hit Database Source	Top Hit Descriptor
12123	24882	38583	1.89	8.7E-02	Z74080.1	NT	S.cerevisiae chromosome IV reading frame ORF YDL012c
12123	24882	38594	1.89	8.7E-02	Z74080.1	NT	S.cerevisiae chromosome IV reading frame ORF YDL012c
12428	25186		1.6	8.7E-02	X17116.1	NT	Human DNA for immunoglobulin alpha heavy chain from a case of alpha heavy chain disease
1257	14282	27256	6.2	8.6E-02	AJ271738.1	NT	Homo sapiens Xq pseudautosomal region; segment 2/2
2256	16270	28295	1.98	8.6E-02	BE408667.1	EST_HUMAN	601304016F1 NIH_MGC_21 Homo sapiens cDNA clone IMAGE:3638643 5'
3202	18257	29176	4.47	8.6E-02	L05468.1	NT	Trichomonas vaginalis beta-tubulin (btub1) gene, complete cds
3658	16701		4.37	8.6E-02	AF153362.1	NT	Dictyostelium discoideum adenylate cyclase (acrA) gene, complete cds
3783	16834		0.68	8.6E-02	U28187.1	NT	Mus musculus long incubation prion protein (Prnp) and prion-like protein (Prnd) genes, complete cds
4512	17537	30421	0.67	8.6E-02	U88179.1	NT	Oryctolagus cuniculus galectin-3 gene, untranslated exon and 5' flanking region
5170	18178		1.09	8.6E-02	L13418.1	NT	Chromatium vinosum tetraene cyclochrome c gene, 3' end, bacterial antylin homologue, flavocytochrome c
5276	18282		1.09	8.6E-02	AB011163.1	NT	Homo sapiens fcaA (complete cds) and flavin subunit, fcaB (3' end)
6213	19287	32520	4.24	8.6E-02	Y10826.1	NT	Homo sapiens mRNA for KIAA0581 protein, partial cds
6510	18574	32628	1.48	8.6E-02	J00440.1	NT	Homo sapiens LCN1b gene
6510	18574	32629	1.48	8.6E-02	J00440.1	NT	Mouse germline IgM chain gene, D region; D-q52, mu switch region (part a)
7834	20781	34159	1.01	8.6E-02	P14616	SWISSPROT	Mouse germline IgM chain gene, D region; D-q52, mu switch region (part a)
8282	21231	34640	1.33	8.6E-02		NT	INSULIN RECEPTOR-RELATED PROTEIN PRECURSOR (IRR) (IR-RELATED RECEPTOR)
8282	21231	34641	1.33	8.6E-02	5730066	NT	Homo sapiens Srf2-related CBP activator protein (SRCAP) mRNA
8405	21374	34782	0.66	8.6E-02	11427428	NT	Homo sapiens Srf2-related CBP activator protein (SRCAP) mRNA
8469	21438		0.81	8.6E-02	U60168.1	NT	Homo sapiens hypothetical protein FLJ11008 (FLJ11008), mRNA
10084	23020	36495	1.04	8.6E-02	AF11170.3	NT	Dictyostelium discoideum proteasome subunit C2 homolog PrtC (prtC) gene, complete cds
10131	23057		1.57	8.6E-02	AW662153.1	EST_HUMAN	Homo sapiens 14q32 Jagged2 gene, complete cds; and unknown gene
10513	23435	36933	0.97	8.6E-02	AF026504.1	NT	h2Dc08.x1 NCI_CGAP_GU1 Homo sapiens cDNA clone IMAGE:2872846 3'
11579	24517	38072	2.47	8.6E-02	BF305608.1	EST_HUMAN	Rattus norvegicus SPA-1 like protein p1294 mRNA, complete cds
11579	24517	38073	2.47	8.6E-02	BF305608.1	EST_HUMAN	601893437F1 NIH_MGC_17 Homo sapiens cDNA clone IMAGE:4139216 5'
11767	23922	37441	9.23	8.6E-02	AE001073.1	NT	601893437F1 NIH_MGC_17 Homo sapiens cDNA clone IMAGE:4139216 5'
							Archaeoglobus fulgidus section 34 of 172 of the complete genome
							Bacillus stearothermophilus BsrF1 methylase (FIM) and BsrF1 restriction endonuclease (FIR) genes, complete cds
11901	24782	38370	1.57	8.6E-02	AF283680.1	NT	cds
2408	15413	28437	2.86	8.5E-02	AE000652.1	NT	Helicobacter pylori 26695 section 130 of 194 of the complete genome
							cg83507.s1 NCI_CGAP_Kid6 Homo sapiens cDNA clone IMAGE:1592917 3' similar to gbX01144 HLA
5762	18946	32028	0.71	8.5E-02	AA985491.1	EST_HUMAN	CLASS II HISTOCOMPATIBILITY ANTIGEN, GAMMA CHAIN PRECURSOR (HUMAN);
5783	18885		1.92	8.5E-02	P08089	SWISSPROT	M PROTEIN, SEROTYPE 6 PRECURSOR
6127	19205	32428	6.34	8.6E-02	AF233885.1	NT	Mus musculus phospholipase C-like protein mRNA, partial cds

Table 4

Single Exon Probes Expressed in Bone Marrow

Probe SEQ ID NO:	Exon SEQ ID NO:	ORF SEQ ID NO:	Expression Signal	Most Similar (Top) Hit BLAST E Value	Top Hit Accession No.	Top Hit Database Source	Top Hit Descriptor
8953	21919	35345	1.02	8.5E-02	8754779	NT	Mus musculus myosin XV (Myo15), mRNA
10198	23121	36807	3.17	8.5E-02	BE833054.1	EST_HUMAN	RC4-OT0037-200700-014-e05 OT0037 Homo sapiens cDNA
10196	23121	36808	3.17	8.5E-02	BE833054.1	EST_HUMAN	RC4-OT0037-200700-014-e05 OT0037 Homo sapiens cDNA
10338	23282	36742	0.44	8.5E-02	A1140618.1	EST_HUMAN	q605g05.x1 Soares_testis_NHTT Homo sapiens cDNA clone IMAGE:1738136 3'
10728	23650	37143	0.56	8.5E-02	X76731.1	NT	V. armodytes gene for armodytoxin C
10851	23771	37270	1	8.5E-02	11418108	NT	Homo sapiens chromosome 22 open reading frame 5 (C22ORF5), mRNA
11490	24433		8.1	8.5E-02	AF155510.1	NT	Homo sapiens heparanase precursor, mRNA, complete cds
11507	24449	37888	4.05	8.5E-02	AB001582.1	NT	Streptococcus mutans gene for glucose-1-phosphate uridylyltransferase, complete cds
12985	26547		5.18	8.5E-02	AA362834.1	EST_HUMAN	EST172738 Ovary II Homo sapiens cDNA 5' end
2676	15901	28690	3.69	8.4E-02	W60330.1	EST_HUMAN	z044e11.11 Soares_fetal_heart_NBH-H19W Homo sapiens cDNA clone IMAGE:343532 5'
4381	17409	30280	1.02	8.4E-02	AF257213.1	NT	Cavia porcellus glycoprotein alpha-subunit mRNA, complete cds
4381	17409	30281	1.02	8.4E-02	AF257213.1	NT	Cavia porcellus glycoprotein alpha-subunit mRNA, complete cds
5235	18243	31115	0.98	8.4E-02	5453817	NT	Homo sapiens nucleobindin 1 (NUCB1), mRNA
5385	18488	31363	8.74	8.4E-02	BE267163.1	EST_HUMAN	601190436F1 NIH_MGC_7 Homo sapiens cDNA clone IMAGE:3534383 5'
6847	19300	33194	1.86	8.4E-02	AK024458.1	NT	Homo sapiens mRNA for FLJ00050 protein, partial cds
8363	21332	34744	7.84	8.4E-02	BE085074.1	EST_HUMAN	CM3-B10780-280400-162-d05 B10780 Homo sapiens cDNA
9194	22160	35588	1.02	8.4E-02	AF218880.1	NT	Homo sapiens atrial natriuretic precursor (ATRNP) gene, exon 2
10727	23649	37142	1.9	8.4E-02	AI735184.1	EST_HUMAN	as88g10.x1 Barstead codon HPLRB7 Homo sapiens cDNA clone IMAGE:2336842 3' similar to TR:O88312
12351	25148	31853	1.46	8.4E-02	R79408.1	EST_HUMAN	O88312 GOB-4. ;
2027	15047	28060	0.87	8.3E-02	5835680	NT	h83h12.11 Soares placenta Nb2HP Homo sapiens cDNA clone IMAGE:145895 5'
2027	15047	28061	0.87	8.3E-02	5835680	NT	h83h12.11 Soares placenta Nb2HP Homo sapiens cDNA clone IMAGE:145895 5'
3606	16651	28568	6.11	8.3E-02	P75334	SWISSPROT	h83h12.11 Soares placenta Nb2HP Homo sapiens cDNA clone IMAGE:145895 5'
3634	16677	28590	0.66	8.3E-02	AI436797.1	EST_HUMAN	h83h12.11 Soares placenta Nb2HP Homo sapiens cDNA clone IMAGE:145895 5'
3634	16677	28591	0.66	8.3E-02	AI436797.1	EST_HUMAN	h83h12.11 Soares placenta Nb2HP Homo sapiens cDNA clone IMAGE:145895 5'
6389	19457	32702	0.78	8.3E-02	AI042338.1	EST_HUMAN	h83h12.11 Soares placenta Nb2HP Homo sapiens cDNA clone IMAGE:145895 5'
6502	19598	32818	2.54	8.3E-02	AF052683.1	NT	w078f11.x1 NC1_CGAP_Kid11 Homo sapiens cDNA clone IMAGE:2461581 3'
8313	21282	34694	3.61	8.3E-02	AF195787.1	NT	Homo sapiens probocardin 43 gene, exon 1
8346	21315		1.19	8.3E-02	AA865285.1	EST_HUMAN	Rattus norvegicus dystrophin-related protein 2 A-form splice variant (Dp2) mRNA, complete cds
8942	21610		1.54	8.3E-02	AA865285.1	EST_HUMAN	og88g08.s1 NC1_CGAP_Kid11 Homo sapiens cDNA clone IMAGE:1455422 3' similar to contigs L1.1 L1 L1
9898	22849	36306	1.58	8.3E-02	AW683503.1	EST_HUMAN	og88g08.s1 NC1_CGAP_Kid11 Homo sapiens cDNA clone IMAGE:1455422 3' similar to contigs L1.1 L1 L1
9909	22730		2.19	8.3E-02	AL161595.2	NT	repetitive element ;
							og88g08.s1 NC1_CGAP_Kid11 Homo sapiens cDNA clone IMAGE:1592779 3'
							og88g08.s1 NC1_CGAP_Kid11 Homo sapiens cDNA clone IMAGE:1592779 3'
							le05h10.x1 Human Pancreatic islets Homo sapiens cDNA 3' similar to TR:Q16332 Q16332 GAMMA
							SUBUNIT OF SODIUM POTASSIUM ATPASE LIKE ;
							Arabidopsis thaliana DNA chromosome 4, contig fragment No. 91

Table 4

Single Exon Probes Expressed in Bone Marrow

Probe SEQ ID NO:	Exon SEQ ID NO:	ORF SEQ ID NO:	Expression Signal	Most Similar (Top) Hit BLAST E Value	Top Hit Accession No.	Top Hit Database Source	Top Hit Descriptor
10706	23628		0.5	8.3E-02	AF020408.1	NT	Dicotyledon discoidium DocA (docA) mRNA, complete cds
12444	25024		1.48	8.3E-02	BE968498.1	EST_HUMAN	601644770F1 NIH_MGC_56 Homo sapiens cDNA clone IMAGE:3928693 5'
1379	14413		7.44	8.2E-02	Y08170.2	NT	Gallus gallus mRNA for OBCAM protein gamma isoform
1494	14527	27499	1.77	8.2E-02	AF167077.2	NT	Canis familiaris glutamate transporter (EAA14) mRNA, complete cds
3089	16147		2.24	8.2E-02	AL163208.2	NT	Homo sapiens chromosome 21 segment HS21C008
3816	16855		1.61	8.2E-02	AL161498.2	NT	Arabidopsis thaliana DNA chromosome 4, contig fragment No. 10
4033	17071	28972	1.22	8.2E-02	AL163206.2	NT	Homo sapiens chromosome 21 segment HS21C008
4313	17342	30222	5.83	8.2E-02	P48960	SWISSPROT	LEUCOCYTE ANTIGEN CD87 PRECURSOR
4313	17342	30223	5.83	8.2E-02	P48960	SWISSPROT	LEUCOCYTE ANTIGEN CD87 PRECURSOR
4313	17342	30224	5.83	8.2E-02	P48960	SWISSPROT	LEUCOCYTE ANTIGEN CD87 PRECURSOR
5103	18118	30982	1.29	8.2E-02	AF240778.1	NT	Mus musculus pepsinogen F (Pepf) mRNA, complete cds
5118	18128	31004	4.13	8.2E-02	U78009.1	NT	Mus musculus zinc transporter (Znt-3) gene, complete cds
5300	18128	31004	0.95	8.2E-02	U78009.1	NT	Mus musculus zinc transporter (Znt-3) gene, complete cds
5408	18511	31389	1.47	8.2E-02	BE897030.1	EST_HUMAN	601439576F1 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:3924523 5'
7221	20243	33677	2.84	8.2E-02	AF309555.1	NT	Bos taurus connective tissue growth factor precursor (CTGF) gene, complete cds
8005	20944		0.6	8.2E-02	AV743341.1	EST_HUMAN	AV743341 CB Homo sapiens cDNA clone CBLANF07 5'
9057	22023		0.45	8.2E-02	U28397.1	NT	Rattus norvegicus plasma membrane Ca2+ ATPase isoform 3 (PMCA3) gene, 5' flanking region
9124	22090	35518	2.75	8.2E-02	AW875128.1	EST_HUMAN	RC2-PT0004-031299-011-c05 P10004 Homo sapiens cDNA
8858	22885	36346	5.43	8.2E-02	X04197.1	NT	Beet necrotic yellow vein virus RNA-2
10121	23047	36526	2.38	8.2E-02	BE254318.1	EST_HUMAN	601115055F1 NIH_MGC_18 Homo sapiens cDNA clone IMAGE:39355698 5'
12450	25209	31831	6.88	8.2E-02	AE002246.2	NT	Chlamydia pneumoniae AR39, section 73 of 94 of the complete genome
12835	25710		5.74	8.2E-02	AF276366.1	NT	Mus musculus epidermal growth factor receptor (Egfr) gene, exons 5 through 28, and complete cds, alternatively spliced
1493	14528	27498	1.05	8.1E-02	AB017138.1	NT	Pseudomonas putida malonate decarboxylase gene cluster (mdcA, mdcB, mdcC, mdcD, mdcE, mdcG, mdcH, mdcI, and mdcM genes), complete cds
5848	18638	32123	1.05	8.1E-02	AE004006.1	NT	Xylella fastidiosa, section 152 of 228 of the complete genome
6516	18579	32636	1.38	8.1E-02	T11532.1	EST_HUMAN	A1484F Heart Homo sapiens cDNA clone A1484
7403	20371		0.81	8.1E-02	AL163279.2	NT	Homo sapiens chromosome 21 segment HS21C079
7835	20782		0.9	8.1E-02	AI692681.1	EST_HUMAN	wc86f08.x1 NCI_CGAP_Lu24 Homo sapiens cDNA clone IMAGE:2336503 3'
8683	21651	35072	0.53	8.1E-02	11428974	NT	Homo sapiens hypothetical protein FLJ10060 (FLJ10060), mRNA
8683	21651	35073	0.53	8.1E-02	11428974	NT	Homo sapiens hypothetical protein FLJ10060 (FLJ10060), mRNA
10272	23197		1.58	8.1E-02	AY005150.1	NT	Homo sapiens extracellular glycoprotein lactrin precursor, gene, complete cds
11827	24710	38294	1.53	8.1E-02	AL163202.2	NT	Homo sapiens chromosome 21 segment HS21C002
6	15830	26026	4.82	8.0E-02	AW954653.1	EST_HUMAN	EST368723 MAGE resequences, MAGEC Homo sapiens cDNA

Table 4

Single Exon Probes Expressed in Bone Marrow

Probe SEQ ID NO.	Exon SEQ ID NO.	ORF SEQ ID NO.	Expression Signal	Most Similar (Top) Hit BLAST E Value	Top Hit Accession No.	Top Hit Database Source	Top Hit Descriptor
1709	15875	27722	13.63	8.0E-02	D26535.1	NT	Human gene for dihydrolipoamide succinyltransferase, complete cds (exon 1-15)
1709	15875	27723	13.63	8.0E-02	D26535.1	NT	Human gene for dihydrolipoamide succinyltransferase, complete cds (exon 1-15)
1919	14943	27839	4.07	8.0E-02	BE067219.1	EST_HUMAN	PM3-BT0347-170200-001-508 BT0347 Homo sapiens cDNA
2384	15392	28417	1.05	8.0E-02	D90815.1	NT	Synechocystis sp. PCC8803 complete genome, 17127, 2137259-2287259
2384	15392	28418	1.05	8.0E-02	D90815.1	NT	Synechocystis sp. PCC8803 complete genome, 17127, 2137259-2287259
2478	15482		4.69	8.0E-02	BF246744.1	EST_HUMAN	601855548F1 NIH_MGC_57 Homo sapiens cDNA clone IMAGE:4075619 5'
2831	14137	27088	0.99	8.0E-02	M28449.1	NT	Dictyosellum discoideum cyclic nucleotide phosphodiesterase gene, complete cds
2911	15989	28892	0.78	8.0E-02	AL445087.1	NT	Thermoplasma acidophilum complete genome; segment 5/5
3630	16870	28172	6.64	8.0E-02	AW966118.1	EST_HUMAN	EST378191 MAGI resequences, MAGI Homo sapiens cDNA
4810	17827	30724	1.7	8.0E-02	AL434202.1	EST_HUMAN	831402.x1 NCI_CGAP_Gas4 Homo sapiens cDNA clone IMAGE:2132114 3'
4818	17835	30733	0.99	8.0E-02	AF116558.1	NT	Arabidopsis thaliana putative transcription factor (HUA2) mRNA, complete cds
4832	17869		7.57	8.0E-02	X72794.1	NT	M.musculus gene for gelatinase B
5832	18022	32105	0.71	8.0E-02	AW981139.1	EST_HUMAN	EST383209 MAGI resequences, MAGI Homo sapiens cDNA
5993	18077	32274	3.28	8.0E-02	AF275948.1	NT	Homo sapiens ABCA1 (ABCA1) gene, complete cds
7368	18077	32274	1.44	8.0E-02	AF275948.1	NT	Homo sapiens ABCA1 (ABCA1) gene, complete cds
8485	21434	34851	3.74	8.0E-02	AL114993.1	NT	Botrytis cinerea strain T4 cDNA library under conditions of nitrogen deprivation
9744	22685	36139	1.21	8.0E-02	X74208.1	NT	H. sapiens AGT gene, Intron 4
9744	22685	36140	1.21	8.0E-02	X74208.1	NT	H. sapiens AGT gene, Intron 4
10519	23441		0.8	8.0E-02	AL163209.2	NT	Homo sapiens chromosome 21 segment HS21C009
11145	24105	37632	2.19	8.0E-02	AF217798.1	NT	Homo sapiens SCG10 like-protein, helicase-like protein NHL, M88, and ADP-ribosylation factor related protein 1 (ARFRP1) genes, complete cds
12493	25230	31798	6.54	8.0E-02	AJ005375.1	NT	Drosophila arena hunchback region
13036	18342		2.08	8.0E-02	4503034	NT	Homo sapiens cAMP responsive element binding protein-like 2 (CREBL2) mRNA
2184	15189	28219	4.36	7.8E-02	BE250008.1	EST_HUMAN	60084319F1 NIH_MGC_15 Homo sapiens cDNA clone IMAGE:2855910 5'
2392	16050	28671	8.43	7.9E-02	AI582028.1	EST_HUMAN	ar88c08.x1 Barstead colon HPLRB7 Homo sapiens cDNA clone IMAGE:2173646 3' similar to gb:Z26876
3864	16003	28608	5.68	7.9E-02	6881044	NT	60S RIBOSOMAL PROTEIN L38 (HUMAN);
3864	16003	28609	5.68	7.9E-02	6881044	NT	Mus musculus colony stimulating factor 1 receptor (Csf1r), mRNA
4722	17742	30633	1.08	7.9E-02	BF348454.1	EST_HUMAN	Mus musculus colony stimulating factor 1 receptor (Csf1r), mRNA
4849	17866		1.49	7.9E-02	AB068019.1	NT	602019770F1 NCI_CGAP_Bim67 Homo sapiens cDNA clone IMAGE:4155401 5'
6855	19008		1.06	7.9E-02	BF368016.1	EST_HUMAN	Arabidopsis thaliana RXW24L mRNA, partial cds
8388	21335	34747	3.28	7.9E-02	U27832.1	NT	RC3-GN0042-310800-024-d11 GN0042 Homo sapiens cDNA
10388	23310	36788	4.89	7.9E-02	AI081644.1	EST_HUMAN	Saccharomyces cerevisiae suppressor of Mif2 Smt4p (SMT4) gene, complete cds ou63605.s1 NCI_CGAP_Br2 Homo sapiens cDNA clone IMAGE:1632465 3' similar to WP:C37A2.2 CE08611;

Table 4

Single Exon Probes Expressed in Bone Marrow

Probe SEQ ID NO:	Exon SEQ ID NO:	ORF SEQ ID NO:	Expression Signal	Most Similar (Top) Hit BLAST E Value	Top Hit Accession No.	Top Hit Database Source	Top Hit Descriptor
10388	23310	36789	4.89	7.8E-02	AJ081644.1	EST_HUMAN	cu53b05.a1 NCI_CGAP_Br2 Homo sapiens cDNA clone IMAGE:1632465 3' similar to WP-C37A2.2 CE08611
1215	14253	27210	1.69	7.8E-02	AJ783275.1	EST_HUMAN	cc59d02.y5 NCI_CGAP_Lu5 Homo sapiens cDNA clone IMAGE:1570467 5' similar to contains L1.13 L1 repetitive element
1215	14253	27211	1.69	7.8E-02	AJ783275.1	EST_HUMAN	cc59d02.y5 NCI_CGAP_Lu5 Homo sapiens cDNA clone IMAGE:1570467 5' similar to contains L1.13 L1 repetitive element
6123	16804		3.25	7.8E-02	BE260048.1	EST_HUMAN	600943055F1 NIH_MGC_15 Homo sapiens cDNA clone IMAGE:2659683 5'
7279	20013	33317	1.29	7.8E-02	U82685.2	NT	Homo sapiens zinc finger protein 92 (ZFP92), expressed-Xq28STS protein (XQ28ORF), and biglycan (BGN) genes, complete cds; and plasma membrane calcium ATPase isoform 3 (PMCA3) gene, partial cds
7279	20013	33318	1.29	7.8E-02	U82685.2	NT	Homo sapiens zinc finger protein 92 (ZFP92), expressed-Xq28STS protein (XQ28ORF), and biglycan (BGN) genes, complete cds; and plasma membrane calcium ATPase isoform 3 (PMCA3) gene, partial cds
9137	22103	35529	1.26	7.8E-02	BE987947.1	EST_HUMAN	601440439F1 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:3925449 5'
9232	22198	35529	0.8	7.8E-02	X79344.1	NT	S.cerevisiae CAT8 gene
9408	22373	35908	0.68	7.8E-02	AF233437.1	NT	Homo sapiens FYVE domain-containing dual specificity protein phosphatase FYVE-DSP1b mRNA, complete cds
9408	22373	35909	0.68	7.8E-02	AF233437.1	NT	Homo sapiens FYVE domain-containing dual specificity protein phosphatase FYVE-DSP1b mRNA, complete cds
9716	22744	36195	1.23	7.8E-02	AA498354.1	EST_HUMAN	nc68b06.r1 NCI_CGAP_P1 Homo sapiens cDNA clone IMAGE:771731
10161	23088	36563	0.65	7.8E-02	Z98124.1	NT	Bacillus subtilis complete genome (section 21 of 21); from 3989281 to 4214814
12114	24984	38585	1.84	7.8E-02	BF025981.1	EST_HUMAN	601688979F1 NIH_MGC_20 Homo sapiens cDNA clone IMAGE:3953083 5'
1399	16857	27402	0.93	7.7E-02	AF181897.1	NT	Homo sapiens WRN (WRN) gene, complete cds
3600	16845		2.05	7.7E-02	AJ238093.1	NT	Homo sapiens partial AF-4 gene, exons 2 to 7 and Alu repeat elements
5622	18718	31877	0.61	7.7E-02	AF062636.1	NT	Gallus gallus collagen type XII alpha-1 (COL12A1) gene, promoter region and partial cds
8241	21210	34615	7.34	7.7E-02	AA402949.1	EST_HUMAN	zu63d11.1 Soares ovary tumor NBH07 Homo sapiens cDNA clone IMAGE:741717 5' similar to TR:G1173905 G1173905 SPLICEOSOME ASSOCIATED PROTEIN. ;
10185	23120	36806	3.76	7.7E-02	P38080	SWISSPROT	PROBABLE SERINE/THREONINE-PROTEIN KINASE YBR059C
10490	23412	36909	0.76	7.7E-02	AJ318662.1	EST_HUMAN	1a80b08.x1 NCI_CGAP_HSC2 Homo sapiens cDNA clone IMAGE:2050359 3' similar to gbt-Z26878 60S RIBOSOMAL PROTEIN L38 (HUMAN);
10490	23412	36910	0.76	7.7E-02	AJ318662.1	EST_HUMAN	1a80b08.x1 NCI_CGAP_HSC2 Homo sapiens cDNA clone IMAGE:2050359 3' similar to gbt-Z26878 60S RIBOSOMAL PROTEIN L38 (HUMAN);
11362	24302	37829	4.66	7.7E-02	11422767	NT	Homo sapiens KIAA0628 gene product (KIAA0628), mRNA
12671	25795		1.91	7.7E-02	11436859	NT	Homo sapiens interferon regulatory factor 7 (IRF7), mRNA

Page 132 of 546
Table 4

Single Exon Probes Expressed in Bone Marrow

Probe SEQ ID NO:	Exon SEQ ID NO:	ORF SEQ ID NO:	Expression Signal	Most Similar (Top) Hit BLAST E Value	Top Hit Accession No.	Top Hit Database Source	Top Hit Descriptor
3308	16447	29373	2.43	7.6E-02	BE514432.1	EST_HUMAN	601316428F1 NIH_MGC_8 Homo sapiens cDNA clone IMAGE:3634903 5'
3419	16487	29387	1	7.6E-02	AA298447.1	EST_HUMAN	EST1112214 Cerebellum II Homo sapiens cDNA 5' end similar to similar to protodactherin 43
6216	18290	32523	0.72	7.6E-02	AI061275.1	EST_HUMAN	an25g02.x1 Gessler Wilms tumor Homo sapiens cDNA clone IMAGE:1689730 3'
6491	19558	32806	0.87	7.6E-02	BE378328.1	EST_HUMAN	601238402F1 NIH_MGC_44 Homo sapiens cDNA clone IMAGE:3608401 5'
9725	22753	36206	1.35	7.6E-02	AI131016.1	NT	Homo sapiens SCL gene locus
10257	23182		1.42	7.6E-02	AL139078.2	NT	Campylobacter jejuni NCTC11168 complete genome; segment 516
10580	23502	36594	0.47	7.6E-02	BE708002.1	EST_HUMAN	RC1-HT0545-020800-017-008 HT0545 Homo sapiens cDNA
10714	23636		0.66	7.6E-02	BE659638.2	EST_HUMAN	601654915R1 NIH_MGC_57 Homo sapiens cDNA clone IMAGE:38338810 3'
10952	23872	37384	0.67	7.6E-02	X92658.1	NT	L. esculentum mRNA for triose phosphate translocator
10952	23872	37385	0.67	7.6E-02	X92658.1	NT	L. esculentum mRNA for triose phosphate translocator
11988	24863	38459	1.9	7.6E-02	AW989845.1	EST_HUMAN	QV3-BN0048-150400-151-e04 BN0048 Homo sapiens cDNA
785	13845	28790	2.89	7.6E-02	5802083	NT	Homo sapiens solute carrier family 8 (neurotransmitter transporter, glycine), member 9 (SLC8A9), mRNA
785	13845	28791	2.89	7.6E-02	5802083	NT	Homo sapiens solute carrier family 8 (neurotransmitter transporter, glycine), member 9 (SLC8A9), mRNA
1935	14959	27855	0.93	7.6E-02	AL163278.2	NT	Homo sapiens chromosome 21 segment HS21C078
4338	17559	30447	0.63	7.6E-02	AB015981.1	NT	Homo sapiens IL-18 gene for interleukin-18, intron 1 and exon 2
5952	19037	32233	0.71	7.6E-02	AI948714.1	EST_HUMAN	wq24h08.x1 NCI_CGAP_Kid11 Homo sapiens cDNA clone IMAGE:2477257 3'
8881	21649	35071	1.01	7.6E-02	AI884387.1	EST_HUMAN	wf32b02.x1 NCI_CGAP_Bm25 Homo sapiens cDNA clone IMAGE:2428491 3' similar to gp.M14328 ALPHA
8952	21819	35239	1.29	7.6E-02	AU116913.1	EST_HUMAN	ENOLASE (HUMAN);
10392	23314		0.45	7.6E-02	BF221730.1	EST_HUMAN	AU116913 HEMBA1 Homo sapiens cDNA clone HEMBA1000284 5'
10858	23779	37279	0.9	7.6E-02	BF208809.1	EST_HUMAN	7c61c05.x1 NCI_CGAP_P28 Homo sapiens cDNA clone IMAGE:3578504 3' similar to contains element
10953	23873	37386	0.68	7.6E-02	X79480.1	NT	MER27 repetitive element;
478	13550	28478	1.24	7.4E-02	AW838547.1	EST_HUMAN	601870205F1 NIH_MGC_19 Homo sapiens cDNA clone IMAGE:4100449 5'
1456	14489		1.08	7.4E-02	AF030027.1	NT	C.fiml DSM 20113 16S rDNA
2585	15586		0.89	7.4E-02	6755069	NT	RC5-L T0054-280100-011-H09 L T0054 Homo sapiens cDNA
3608	16653	29571	0.78	7.4E-02	AI807885.1	EST_HUMAN	Equine herpesvirus 4 strain NS80587, complete genome
4733	17763	30646	1.09	7.4E-02	L78810.1	EST_HUMAN	Mus musculus paired-like homeodomain transcription factor 1 (Pitx1), mRNA
4833	17850	30750	3.24	7.4E-02	6978442	NT	wf43h01.x1 Soares_NFL_T_GBC_S1 Homo sapiens cDNA clone IMAGE:2358385 3'
4979	17894	30883	2.52	7.4E-02	6678482	NT	Homo sapiens ADP/ATP carrier protein (ANT-2) gene, complete cds
6843	19701		1.8	7.4E-02	R17477.1	EST_HUMAN	Rattus norvegicus Activin receptor like kinase 1 (Acvrl1), mRNA
7707	20684	34031	0.79	7.4E-02	AA605132.1	EST_HUMAN	Mus musculus ubiquitin c-terminal hydrolase related polypeptide (Uchrrp), mRNA
							y914906.r1 Soares infant brain 1N1B Homo sapiens cDNA clone IMAGE:32339 5'
							nc71402.e1 NCI_CGAP_AA1 Homo sapiens cDNA clone IMAGE:1112259 3'

Table 4

Single Exon Probes Expressed In Bone Marrow

Probe SEQ ID NO:	Exon SEQ ID NO:	ORF SEQ ID NO:	Expression Signal	Most Similar (Top) Hit BLAST E Value	Top Hit Accession No.	Top Hit Database Source	Top Hit Descriptor
8233	21202	34608	1.2	7.4E-02	BE880112.1	EST_HUMAN	601483368F1 NIH_MGC_68 Homo sapiens cDNA clone IMAGE:3885264 5'
8848	21813	35233	1.01	7.4E-02	U58089.1	NT	Human periodic tyrocytic protein 2 (PWP2) gene, exons 15 to 21, and complete cds
9521	22484	35830	1.02	7.4E-02	AW628605.1	EST_HUMAN	h167d11.y1 NCI_CGAP_GU1 Homo sapiens cDNA clone IMAGE:2867861 5' similar to SW:SCA2_HUMAN
9521	22484	35831	1.02	7.4E-02	AW628605.1	EST_HUMAN	O15127 SECRETORY CARRIER-ASSOCIATED MEMBRANE PROTEIN 2.:
9784	21117	34516	0.48	7.4E-02	AI872839.1	EST_HUMAN	h167d11.y1 NCI_CGAP_GU1 Homo sapiens cDNA clone IMAGE:2867861 5' similar to SW:SCA2_HUMAN
9784	21117	34517	0.48	7.4E-02	AI872839.1	EST_HUMAN	O15127 SECRETORY CARRIER-ASSOCIATED MEMBRANE PROTEIN 2.:
10173	23098	36578	1.07	7.4E-02	U62293.1	NT	h167d11.y1 Soares_Dickgraefe_colon_NHCD Homo sapiens cDNA clone IMAGE:2846819 3'
10301	23228	38708	0.53	7.4E-02	BF512678.1	EST_HUMAN	Human LIM-kinase1 and alternatively spliced LIM-kinase1 (LIMK1) gene, complete cds
12408	25181		1.47	7.4E-02	11525883	NT	U-H-BW1-ang-g-08-Q-UJ.s1 NCI_CGAP_Sub7 Homo sapiens cDNA clone IMAGE:3068898 3'
12665	25903		2.61	7.4E-02	AW379491.1	EST_HUMAN	Homo sapiens histone deacetylase 5 (NY-CO-9), mRNA
489	13542	28468	1.3	7.3E-02	BE964901.2	EST_HUMAN	CM4-HT0243-081189-037-d11 HT0243 Homo sapiens cDNA
489	13542	28469	1.3	7.3E-02	BE964901.2	EST_HUMAN	601668738R1 NIH_MGC_68 Homo sapiens cDNA clone IMAGE:3886209 3'
685	13748	28874	5.48	7.3E-02	AE001789.1	NT	601668738R1 NIH_MGC_68 Homo sapiens cDNA clone IMAGE:3886209 3'
1477	15868	27486	3.94	7.3E-02	AW900281.1	EST_HUMAN	Thermotoga maritima section 101 of 136 of the complete genome
1862	15879		15.78	7.3E-02	AL163302.2	NT	CM0-NN1004-130300-284-g08 NN1004 Homo sapiens cDNA
							Homo sapiens chromosome 21 segment HS21C102
							Human germline T-cell receptor beta chain Dopamine-beta-hydroxylase-like, TRY1, TRY2, TRY3,
							TCRBV27S1P, TCRBV22S1A2N1T, TCRBV6S1A1T, TCRBV7S1A1N2T, TCRBV6S1A1T, TCRBV13S3,
							TCRBV6S7P, TCRBV7S3A2T, TCRBV13S2A1T, TCRBV6S2A2PT, TCRBV7S2A1N4T,
3785	18828		0.72	7.3E-02	U68059.1	NT	TCRBV13S9N13S>
5032	18046		1.06	7.3E-02	U12283.1	NT	Mus musculus transcription factor USF2 (USF2) gene, exons 8-10 and complete cds
6595	19555	32827	1.04	7.3E-02	AA779877.1	EST_HUMAN	z24e02.s1 Soares_fetal_liver_spleen_1NFLS_S1 Homo sapiens cDNA clone IMAGE:451178 3' similar to
7706	20663	34029	2.47	7.3E-02	P05143	SWISSPROT	gb1D2428 26S PROTEASE SUBUNIT 4 (HUMAN);
7706	20663	34030	2.47	7.3E-02	P05143	SWISSPROT	PROLINE-RICH PROTEIN MP-3
8508	21476		1.2	7.3E-02	7682107	NT	PROLINE-RICH PROTEIN MP-3
9665	22827		1.39	7.3E-02	AB011090.1	NT	Homo sapiens KIAA0424 protein (KIAA0424), mRNA
							Homo sapiens mRNA for KIAA0518 protein, partial cds
11552	19555	32827	1.89	7.3E-02	AA779877.1	EST_HUMAN	z24e02.s1 Soares_fetal_liver_spleen_1NFLS_S1 Homo sapiens cDNA clone IMAGE:451178 3' similar to
12825	25510		1.33	7.3E-02	Z73597.1	NT	gb1D2428 26S PROTEASE SUBUNIT 4 (HUMAN);
							S.cerevisiae chromosome XVI reading frame ORF YPL241c
120	13229	26158	0.7	7.2E-02	AE000982.1	NT	Methanobacterium thermoautotrophicum from bases 1028155 to 1038934 (section 88 of 148) of the complete genome

Page 134 of 546
Table 4

Single Exon Probes Expressed in Bone Marrow

Probe SEQ ID NO:	Exon SEQ ID NO:	ORF SEQ ID NO:	Expression Signal	Most Similar (Top) Hit BLAST E Value	Top Hit Accession No.	Top Hit Database Source	Top Hit Descriptor
120	13229	26157	0.7	7.2E-02	AE000882.1	NT	Methanobacterium thermoautotrophicum from bases 1029155 to 1039934 (section 88 of 148) of the complete genome
1472	14505	27478	2.77	7.2E-02	AL163301.2	NT	Homo sapiens chromosome 21 segment HS21C101
1472	14505	27479	2.77	7.2E-02	AL163301.2	NT	Homo sapiens chromosome 21 segment HS21C101
2558	15558		2.98	7.2E-02	U14794.1	NT	Human Immunodeficiency virus type 1 isolate 28 reverse transcriptase (pol) gene, internal fragment, partial cds
3900	16940	29851	0.74	7.2E-02	AW288322.1	EST_HUMAN	U1H-BW0-aj1-a-05-Q-JL1 NCI_CGAP_Sub8 Homo sapiens cDNA clone IMAGE:2732049 3'
4373	17400	30280	4.57	7.2E-02	BF572307.1	EST_HUMAN	602077757F1 NIH_MGC_62 Homo sapiens cDNA clone IMAGE:4251950 5'
5360	18465	31336	2.81	7.2E-02	U67531.1	NT	Methanococcus jannaschii section 73 of 150 of the complete genome
5361	18466	31337	9.14	7.2E-02	P11120	SWISSPROT	CALMODULIN
6239	18812		0.73	7.2E-02	BF217598.1	EST_HUMAN	601883905F1 NIH_MGC_57 Homo sapiens cDNA clone IMAGE:4086224 5'
7374	20344	33668	1.76	7.2E-02	BF218088.1	EST_HUMAN	601883508F1 NIH_MGC_57 Homo sapiens cDNA clone IMAGE:4086710 5'
7391	20360	33712	0.6	7.2E-02	AF221128.1	NT	Streptococcus pneumoniae putative response regulator (zmpR), putative histidine kinase (zmpS), and putative zinc metalloprotease (zmpB) genes, complete cds
7417	20384		1.74	7.2E-02	8534897	NT	Strongylocentrotus purpuratus mitochondrion, complete genome
8529	21497	34912	0.69	7.2E-02	P05143	SWISSPROT	PROLINE-RICH PROTEIN MP-3
8529	21497	34913	0.69	7.2E-02	P05143	SWISSPROT	PROLINE-RICH PROTEIN MP-3
9419	22384		0.62	7.2E-02	Y17217.1	NT	Lactococcus lactis cspE gene
9934	22861		0.63	7.2E-02	X16340.1	NT	Human gene for sex hormone-binding globulin (SHBG)
9970	22897	36380	1.98	7.2E-02	AV172452.1	EST_HUMAN	AV172452 DCA Homo sapiens cDNA clone DCAAUG01 5'
10117	23043	36523	3.63	7.2E-02	L14591.1	NT	Homo sapiens plasma membrane calcium ATPase isoform 1 (ATP2B1) gene, alternative splice products, partial cds
10274	23199	36668	1.17	7.2E-02	BF125399.1	EST_HUMAN	601763523F1 NIH_MGC_20 Homo sapiens cDNA clone IMAGE:4026436 5'
10362	23285	36782	2.29	7.2E-02	AW873187.1	EST_HUMAN	h24F11.x1 NCI_CGAP_Adr1 Homo sapiens cDNA clone IMAGE:3120333 3' similar to TR:Q92340 Q92340 ATYPICAL PKC SPECIFIC BINDING PROTEIN ;
10552	23474	36869	0.54	7.2E-02	AA768204.1	EST_HUMAN	aa62c07.s1 NCI_CGAP_GC81 Homo sapiens cDNA clone IMAGE:1316844 3'
10717	23639	37132	2.17	7.2E-02	U82895.2	NT	Homo sapiens zinc finger protein 92 (ZFP92), expressed-Xq28STS protein (XQ28ORF), and biglycan (BGN) genes, complete cds; and plasma membrane calcium ATPase isoform 3 (PMCA3) gene, partial cds
10840	23760	37280	5.41	7.2E-02	BE585003.1	EST_HUMAN	601343928F1 NIH_MGC_53 Homo sapiens cDNA clone IMAGE:3685651 5'
10894	23784		3.24	7.2E-02	BE539214.1	EST_HUMAN	601065194F1 NIH_MGC_10 Homo sapiens cDNA clone IMAGE:3451559 5'
11258	24210	37733	4.17	7.2E-02	AF049874.1	NT	Rattus norvegicus bHLH transcription factor Mist1 (Mist1) gene, complete cds
12311	25119	31842	1.58	7.2E-02	AA773698.1	EST_HUMAN	af81a04.r1 Soares_NH-IMPu_S1 Homo sapiens cDNA clone IMAGE:1048398 5'
12350	25145		5.13	7.2E-02	AJ230786.1	EST_HUMAN	AJ230786 Homo sapiens library (Seranski P) Homo sapiens cDNA clone PS13D5 3'

Table 4

Single Exon Probes Expressed in Bone Marrow

Probe SEQ ID NO:	Exon SEQ ID NO:	ORF SEQ ID NO:	Expression Signal	Most Similar (Top) Hit BLAST E Value	Top Hit Accession No.	Top Hit Database Source	Top Hit Descriptor
12469	25218		3.99	7.2E-02	U82828.1	NT	Homo sapiens alarida telanglectasia (ATM) gene, complete cds
12484	25754		8.41	7.2E-02	AW900982.1	EST_HUMAN	GM4-NN1009-200300-116-c11 NN1009 Homo sapiens cDNA
12880	25933		1.45	7.2E-02	AF020439.1	NT	Homo sapiens ATP-citrate lyase gene, intron 3
12884	25533		1.85	7.2E-02	AA401779.1	EST_HUMAN	z557c12.r1 Soares_testis_NHT Homo sapiens cDNA clone IMAGE:726454 5'
1920	14944	27940	1.83	7.1E-02	L02280.1	NT	Human Immunodeficiency virus type 1 (D8) proviral structural capsid protein (gag) gene, partial cds
2300	15312	26332	6.39	7.1E-02	BF208802.1	EST_HUMAN	601872281F1 NIH_MGC_53 Homo sapiens cDNA clone IMAGE:4092981 5'
8239	21208	34612	0.87	7.1E-02	A1125284.1	EST_HUMAN	q892x10.x1 Soares_testis_NHT Homo sapiens cDNA clone IMAGE:1736922 3'
12193	25038		8.33	7.1E-02	BE304764.1	EST_HUMAN	601143974F1 NIH_MGC_15 Homo sapiens cDNA clone IMAGE:3051234 5'
529	19800	28518	1.42	7.0E-02	Q07092	SWISSPROT	COLLAGEN ALPHA 1(XVI) CHAIN PRECURSOR
1499	14632		1.82	7.0E-02	X68677.1	NT	Martellia Ntcut-1 gene
1776	14805	27781	1.18	7.0E-02	AA056343.1	EST_HUMAN	z69304.s1 Stragene colon (#637204) Homo sapiens cDNA clone IMAGE:509599 3'
3042	16100	29015	1.76	7.0E-02	AW138152.1	EST_HUMAN	U1H-B11-acy-c-07-0-U1.s1 NCI_CGAP_Sub3 Homo sapiens cDNA clone IMAGE:2716020 3'
3913	16953	28884	0.98	7.0E-02	AA815438.1	EST_HUMAN	af65a12.s1 Soares_testis_NHT Homo sapiens cDNA clone 1375678 3' similar to gb:K03002 60S
4074	17110	30004	1.05	7.0E-02	BE070264.1	EST_HUMAN	RIBOSOMAL PROTEIN L32 (HUMAN);
4164	17195		1.24	7.0E-02	AW752982.1	EST_HUMAN	QV4-BT0407-280100-060-e10 BT0407 Homo sapiens cDNA
4242	17271	30155	1.28	7.0E-02	AF077821.1	NT	CM0-UM0001-060300-270-e12 UM0001 Homo sapiens cDNA
4968	17981	30871	8.41	7.0E-02	BF381987.1	EST_HUMAN	Caris familiaris inducible nitric oxide synthase mRNA, complete cds
5451	18553		1.03	7.0E-02	Y09143.2	NT	601816291F1 NIH_MGC_58 Homo sapiens cDNA clone IMAGE:4050071 5'
7630	20590	33953	0.9	7.0E-02	AV689285.1	EST_HUMAN	Lumbricus rubellus mRNA for cyclophilin B
7894	20808	34189	0.8	7.0E-02	Y19187.1	NT	AV689285 GK Homo sapiens cDNA clone GKCCAE06 5'
9454	22418	35858	1.2	7.0E-02	6828113	NT	Gallus gallus mRNA for partial accorin, XL spliced variant (acc gene)
8959	22893	36345	1.21	7.0E-02	K02901.1	NT	African swine fever virus, complete genome
10312	23236	36718	0.76	7.0E-02	U27286.1	NT	Rat Ig gamma epsilon H-chain gene C-region, 3' end
11700	24685	38242	2.23	7.0E-02	AA742495.1	EST_HUMAN	Human myosin binding protein H (MyBP-H) gene, complete cds
12939	25519	31710	1.68	7.0E-02	11421638	NT	af68a06.s1 Soares_NFL_T_GBC_S1 Homo sapiens cDNA clone IMAGE:1327184 3' similar to gb:U14637
515	13586	26504	11.2	6.9E-02	AL163210.2	NT	TIGHT JUNCTION PROTEIN ZO-1 (HUMAN);
515	13586	26505	11.2	6.9E-02	AL163210.2	NT	Homo sapiens hypothetical protein FLJ20116 (FLJ20116), mRNA
						NT	Homo sapiens chromosome 21 segment HS21C010
						NT	Homo sapiens chromosome 21 segment HS21C010
1336	14370		1.68	6.9E-02	4507968	NT	Homo sapiens regulator of Gs-selective protein signaling (ZGAP1) mRNA, and translated products
3807	18947	29764	1.37	6.9E-02	Q06384	SWISSPROT	28S PROTEASOME REGULATORY SUBUNIT S3 (NUCLEAR ANTIGEN 21D7)
3807	18947	29755	1.37	6.9E-02	Q06384	SWISSPROT	28S PROTEASOME REGULATORY SUBUNIT S3 (NUCLEAR ANTIGEN 21D7)